

Die Größe der effektiven Ionisierungsspannung in der Quecksilberhochdruckentladung

Von R. Borchert

(Mit 5 Abbildungen)

Inhaltsübersicht

Die durchgeführten Betrachtungen entsprechen den tatsächlichen Vorgängen im Bogen nur in guter Näherung; auch ist die Genauigkeit der Intensitäts- und Leitfähigkeitsmessungen nicht groß genug, um die Ionisierungsspannung in der Hg-Hochdruckentladung auf 0,1 Volt genau anzugeben. Das Ergebnis der Messungen kann aber dahin gedeutet werden, daß die Herabsetzung der Ionisierungsgrenze in normalen Hochdruckentladungen bis zu Drucken von 20–30 atm nur etwa 0,2 Volt beträgt und daß eine Herabsetzung bis zu 0,5 Volt erst bei Leistungskonzentration von $> 100 \text{ kW/cm}^2$ erreicht wird, wie sie in den kugelförmigen Höchstdrucklampen mit über 30 atm Druck vorliegen.

Auf die Folgerungen, die hieraus für die Frage nach der Entstehung des Kontinuums und der Größe ihrer mittleren „Anregungsspannung“ gezogen werden müssen, wird in einer weiteren Arbeit eingegangen.

Einleitung: Annahme der Herabsetzung der Ionisierungsgrenze auf Grund der Unsöld'schen Gleichung für die kontinuierliche Strahlung

Für die Quecksilberhochdruckentladung ist von Unsöld¹⁾ eine quantitative Beziehung für die Intensität der kontinuierlichen Strahlung aufgestellt worden. Er weist darauf hin, daß neben den diskret-kontinuierlichen und den kontinuierlich-kontinuierlichen Übergängen bei hohen Drucken auch die stark verbreiterten Terme dicht unterhalb der Ionisierungsgrenze noch einen Beitrag zur kontinuierlichen Absorption liefern. Um diesem Beitrag zur kontinuierlichen Absorption Rechnung zu tragen, integriert Unsöld von $-\infty$ bis $\frac{U_i + \Delta U}{kT}$, wobei U_i die Ionisierungsspannung des isolierten Atoms und ΔU der Bereich der verschmierten Terme bedeuten. Für die gesamte Emission pro cm^2 , Frequenzeinheit und Raumwinkel 4π erhält er somit:

$$4\pi \epsilon_r = \text{Const} \cdot p \cdot e^{-\frac{e(U_i + \Delta U)}{kT}} \quad (1)$$

Ein Vergleich von experimentell gemessenen Intensitäten mit Werten, die sich aus der obigen Gleichung errechnen, ergibt gute Übereinstimmung, sofern man für $(U_i + \Delta U)$ Werte von 9,5 Volt annimmt²⁾. Aus dieser Tatsache wurde die Folgerung

¹⁾ A. Unsöld, Ann. Physik **33**, 607 (1938).

²⁾ P. Schulz, Z. Physik **119**, 167 (1942).

gezogen, daß die Ionisierungsgrenze in der Quecksilberhochdruckentladung um den Wert ΔU herunterrückt; der Wert ($U_i - \Delta U$) wird als effektive Ionisierungsspannung bezeichnet. Da in der Quecksilberhochdruckentladung noch Linien mit der Anregungsspannung von 9,6 Volt festgestellt wurden³⁾, nimmt man für die Ionisierungsspannung den Wert von 9,7 Volt als den wahrscheinlich richtigen an. Das würde bedeuten, daß die Elektronen der Terme bis herunter zu 9,7 Volt nicht mehr an ein bestimmtes Atom gebunden sind, sondern dem elektrischen Feld unterliegen und sich an der Leitfähigkeit beteiligen.

Theoretische Deutung der Zunahme der Elektronenkonzentration aus Elektronen der verschmierten Terme

Rompe und Steenbeck⁴⁾ haben versucht, die eintretende Verringerung der Ionisierungsspannung durch die im Plasma wirkenden Mikrofelder zu erklären und quantitativ abzuschätzen. Um ein Elektron aus dem Atomverband vom Ion zu entfernen, hätten wir im Vakuum die Kraft $\frac{e^2}{r}$ zu überwinden. Im Plasma ist die Kraft durch die sich ausbildende Ladungswolke verkleinert. Dadurch ergibt sich auch eine Verringerung der zu leistenden Ionisierungsarbeit. Die Differenz der Ionisierungsspannung im Vakuum und im Plasma ist gleich der Summe der „Austrittsarbeiten“ von Elektron und Ion aus dem Plasma. Für ein quasi neutrales Plasma ist die Austrittsarbeit für ein Elektron wie für ein Ion $\frac{e^2}{D}$. D bedeutet den Abstand, außerhalb dessen die Feldstärke des zentralen Ions bereits vollständig abgeschirmt ist. Die sich so ergebende Verminderung der Ionisierungsarbeit errechnet sich zu $\sim 5 \cdot 10^{-2}$.

Man kommt zu einer noch größeren Erniedrigung um etwa 0,2 bis 0,3 Volt, wenn man nach Rompe und Steenbeck ein Elektron als praktisch frei ansieht für den Fall, daß die Feldstärke des Zentralfeldes von derselben Größe wird wie das Mikrofeld. Der lineare Mittelwert der absoluten Beträge der Mikrofeldstärke ist

$$|\bar{E}| = A \cdot e \cdot N^{\frac{1}{2}}. \quad (2)$$

Es gilt also die Beziehung:

$$A \cdot e \cdot N e^{\frac{1}{2}} = \frac{e}{r^2}. \quad (3)$$

Für A werden Werte von 12,2, 13,7 und 20 gefunden. r ist der Abstand des Leuchtelektrons vom Atomrest. Setzt man für den Radius einer Bohrschen Bahn $r_a = 0,528 \cdot n^2 \cdot 10^{-8}$ ein, wo n die Hauptquantenzahl des H -Atoms ist, so findet man, daß alle Elektronenzustände mit Hauptquantenzahlen größer als

$$n = 0,7 \cdot 10^4 \cdot N^{-\frac{1}{2}} \quad (4)$$

als praktisch frei zu betrachten sind. Im Quecksilber-Plasma, wo die Elektronenkonzentration etwa 10^{18} pro cm^3 ist, wären demnach Terme mit einer Laufzahl > 7 nicht mehr vorhanden. Die Ionisierungsspannung U_i wäre 0,2 bis 0,3 Volt geringer als die des isolierten Atoms; also 10,2–10,1 Volt.

³⁾ P. Schulz u. W. Weizel, Z. Physik **122**, 697 (1944).

⁴⁾ R. Rompe u. M. Steenbeck, Ergebn. exakt. Naturw. **18**, 257 (1939).

Experimentelle Bestätigung der Zunahme der freien Elektronen aus den verschmierten Termen durch Temperatur-Betrachtung

Eine weitere Verminderung ist theoretisch nicht recht einzusehen. Eine Herabsetzung bis 9,7 Volt setzt voraus, daß wesentlich mehr „Elektronen“ frei gemacht sind und an der Leitfähigkeit teilnehmen, als sich aus der im Plasma herrschenden Temperatur ergibt. Näherungsweise kann man den Bogen als zylinderförmigen Kanal vom Radius R ansetzen, in dem die Temperatur überall konstant ist. Für das Plasma einer Lichtbogensäule läßt sich dann eine einfache Beziehung zwischen Leitfähigkeit und Elektronenkonzentration aufstellen, wenn man den Anteil des Ionenstroms am Gesamtstrom vernachlässigt und berücksichtigt, daß in Richtung der Achse der Entladung und außerhalb der Nähe der Elektroden ist:

$$\sigma = \frac{|i|}{|\mathcal{E}|} = e b_e \cdot n_e. \quad (5)$$

Die Elektronenkonzentration n_e kann aus der Messung der Breite der 3D-Terme bestimmt werden, da diese Verbreiterung auf Elektronen bzw. Ionenstoß zurückzuführen ist. Die durch Elektronenstoß verursachte Verbreiterung ergibt sich zu:

$$\delta \sim T^{\frac{1}{2}} n_e. \quad (6)$$

Aus Gl. (5) erhalten wir für konstanten Druck:

$$\sigma \sim T^{-\frac{1}{2}} n_e, \quad (5a)$$

wenn wir $b_e \sim T^{-\frac{1}{2}}$ ansetzen und die Herabsetzung der Elektronenbeweglichkeit durch den Einfluß der positiven Ionen vernachlässigen⁵⁾.

Mit Gl. (5a) und (6) erhält man

$$T \sim \left(\frac{\delta}{\sigma} \right)^{1,5}. \quad (7)$$

Wir haben damit eine einfache Gleichung für die Temperaturänderung, die sich bei einer Änderung der Leitfähigkeit aus der Erhöhung der Elektronenkonzentration ergibt. Die sich hieraus errechnenden Temperaturwerte werden nur dann mit den im Bogen herrschenden Temperaturen übereinstimmen, wenn die verschmierten Terme keinen Beitrag zur Elektronenkonzentration liefern.

Die im Bogen vorliegenden Temperaturen lassen sich mit ausreichender Genauigkeit aus der Gesamtstrahlung berechnen, die proportional der aufgenommenen

Leistung $L = \frac{I}{\pi R^2} \cdot \mathcal{E}$ Watt/cm² ist. Für die gesamte Strahlung gilt:

$$S \sim N_e \frac{e U_m}{k T} \sim \frac{p}{k T} e \frac{e U_m}{k T}, \quad (8)$$

wobei N die Anzahl der Atome und U_m die mittlere Anregungsspannung bedeuten. U_m kann für den Fall, daß die Strahlung im wesentlichen durch Temperaturanregung entsteht, zu 8 Volt angenommen werden.

Bei bekannter Leistungsänderung $\frac{L_2}{L_1}$ ergibt sich für konstanten Druck für das Verhältnis der Temperaturen

$$\frac{1}{T_1} - \frac{1}{T_2} = \frac{k}{e U_m} \ln \left(\frac{L_2}{L_1} \cdot \frac{T_2}{T_1} \right). \quad (9)$$

⁵⁾ P. Schulz, Ann. Physik, 6. Folge 1, 318 (1947).

Der Ausdruck $\ln \frac{T_2}{T_1}$ kann wegen der geringen Änderung von T vernachlässigt werden.

$$\frac{T_2}{T_1} = 1 + \frac{T_2 \cdot k}{e U_m} \ln \frac{L_2}{L_1} \quad (9a)$$

Das Temperaturverhältnis läßt sich bestimmen, wenn eine der beiden Temperaturen bekannt ist, jedoch braucht dieser Wert nur größenordnungsmäßig richtig zu sein.

In Abb. 1 ist die Breite der 3 D-Terme in Abhängigkeit von der Leitfähigkeit aufgetragen, wie sie sich aus Messungen von Rompe und Schulz an Bögen mit einem Druck von 30 atm ergibt. Da bei den Aufnahmen zur Bestimmung der Linienbreiten die Mitte des Entladungsbogens auf den Spalt des Gitterspektrographen abgebildet wurde, die Temperatur aber von der Bogenachse zum Rande etwa nach der Funktion $y = a e^{-b x^2}$ (Glockenkurve) abfällt, erhält man Elektronendichten, die etwas höher liegen als die über den ganzen Querschnitt gemittelten Werte. Wenn man obige Funktion über den Verlauf der Elektronendichten bzw. der Temperatur annimmt, ergibt sich für einen Kanaldurchmesser von 2 mm und eine Spaltbreite von 0,02 mm ein Fehler von nur etwa 3%. Das liegt innerhalb der Meßgenauigkeit, die zwischen 10–20% liegt.

Bei den zur Messung benutzten Bögen hoher Leistungskonzentration handelt es sich um elektrodenstabilisierte Bögen, auf die die Beziehungen der zylindrischen Bogensäule ebenfalls in guter Näherung angewandt werden können, wie Rompe und Weizel⁶⁾ gezeigt haben. In die Bestimmung der Leitfähigkeitsänderung bei Stromstärke-

Erhöhung geht die Verbreiterung des Bogenkanals quadratisch ein. Das Verhalten des für die Leitfähigkeit maßgebenden Kanaldurchmessers muß daher möglichst genau festgestellt werden.

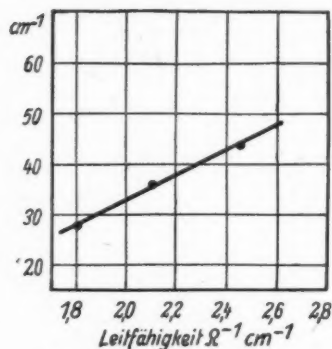


Abb. 1. Breite der 3 D-Terme

Erhöhung geht die Verbreiterung des Bogenkanals quadratisch ein. Das Verhalten des für die Leitfähigkeit maßgebenden Kanaldurchmessers muß daher möglichst genau festgestellt werden.

Rompe und Schulz⁷⁾ haben die Kanalbreite aus der Leuchtdichteverteilung quer zur Bogenachse bestimmt. Streng genommen entspricht der hieraus gewonnene Querschnitt nicht dem für die Leitfähigkeit maßgeblichen, da in die Leitfähigkeitsbeziehung nicht die Strahlung, sondern die Elektronendichte eingeht und

die Temperatur in beide nach dem Boltzmann-Faktor $e^{-\frac{e U_m}{k T}}$ bzw. $e^{-\frac{e U_i}{k T}}$ verschieden eingeht entsprechend dem Unterschied zwischen $U_m = 8$ Volt und $U_i = \sim 10$ Volt⁸⁾.

Für die Bestimmung der Leuchtdichteverteilung gilt noch Folgendes: Die gemessenen Intensitätswerte in den verschiedenen Abständen von der Achse der Entladung stellen Mittelwerte über die Verteilung in parallel zur Achse gelegenen Ebenen dar. Man muß also die gemessene Intensitätsverteilung quer zum Kanaldurchmesser umrechnen. Dieses Verfahren ist nur dann möglich, wenn die Reabsorption vernachlässigbar ist. Für die Gesamtstrahlung einer Hg-Höchstdruckent-

⁶⁾ R. Rompe, W. Thouret u. W. Weizel, Z. Physik **122**, 1 (1944).

Entladung ist das zutreffend. Rompe und Schulz⁷⁾ haben diese Umrechnung durchgeführt und festgestellt, daß gemessene und wahre Intensitätsverteilung für den vorliegenden Fall praktisch gleich sind. In Tabelle 1 sind neben diesen Ergebnissen

Tabelle 1
Halbwertsbreite, Feldstärke, Leistungskonzentration bei Hg-Hochdruckbögen

p	J	U	E	Kanal- Radius	Leistung pro cm^2	Leit- fähigkeit	Breite der 3 D-Terme
atm	Amp.	Volt	V/cm	10^{-2} cm	kW/cm^2	$\Omega^{-1} \text{cm}^{-1}$	cm^{-1}
30	6	73,5	130	9	32,5	1,8	27,8
	8	76,8	137	9,4	39,5	2,1	35,9
	10	80,5	145	9,5	51,2	2,4	44,0
35	3,5	70	129	9,1	17,5	1,05	
	5	74,5	138,5	9,2	26	1,4	
	6,5	77,3	144,5	9,2	35,5	1,7	
	8	80,3	151,5	9,3	44,5	1,95	
45	3,25	74,5	146,5	6,2	39	1,85	
	5,25	80,5	160	6,2	69,5	2,7	
	5,75	81,5	162,5	6,3	75,5	2,85	
	6,3	82,5	164,5	6,4	80	2,95	
	7,5	84	168,5	6,4	98	3,45	
60	2	53,3	173	4,3	59,5	2,0	
	3	55,9	184	4,4	90,5	2,65	
	4	57,3	189	4,5	119	3,35	
	5	59,5	198	4,6	150	3,8	

Meßwerte an Bögen noch höherer Leistungskonzentration aufgeführt. Sie zeigen, daß der für die Strahlung maßgebliche Kanaldurchmesser bei elektrodenstabilisierten Bögen hoher Konzentration praktisch unabhängig von der Leistung ist. Dasselbe kann mit ausreichender Genauigkeit von dem für die Leitfähigkeit maßgebenden Durchmesser angenommen werden.

Aus dem oben ausführlich Gesagten ersieht man, daß es sich bei einem Vergleich der sich aus Gl. (7) und (9a) ergebenden Werte für die Temperaturerhöhung durch Leistungserhöhung bei konstantem Druck nur um eine Abschätzung handeln kann. Die aus beiden Beziehungen

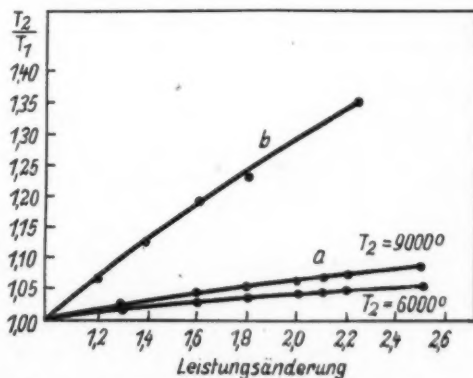


Abb. 2. Temperaturänderung als Funktion der Leistungsänderung. a) Berechnet aus der Beziehung für die Gesamtstrahlung, b) berechnet aus der Elektronenkonzentration, die aus der Breite der 3 D-Terme bestimmt wurde

⁷⁾ R. Rompe u. P. Schulz, Z. Physik 112, 691 (1939).

errechneten Werte sind in Abb. 2 als Funktion der Leistungsänderung aufgetragen. Es ist ein deutlicher Unterschied vorhanden: Die aus der Elektronenkonzentration geschätzte Temperaturerhöhung ist weit größer als die, die sich aus der Gesamtstrahlung berechnet. Dieser Unterschied ist reell und liegt außerhalb der Ungenauigkeit der Abschätzung. Daraus kann der Schluß gezogen werden, daß neben der Erhöhung der Elektronenkonzentration durch den Anstieg der Temperatur noch ein zusätzlicher Anteil an Elektronen hinzukommt, der aus den oberen verschmierten Termen kommen muß. Die Ionisierungsspannung wird also — wie angenommen — kleiner. Die Abschätzung ist aber nicht so exakt, daß man daraus den Wert der effektiven Ionisierungsspannung angeben könnte.

Bestimmung der effektiven Ionisierungsspannung aus der Linienstrahlung

Eine weitere Beziehung für die Elektronenkonzentration ist uns durch die Saha-Gleichung gegeben. Für konstanten Druck ist:

$$n_e \sim T^{\frac{1}{2}} e^{-\frac{e U_i}{k T}}. \quad (10)$$

Unter Benutzung von Gl. (5a) erhält man:

$$e^{-\frac{e U_i}{2 k T}} \sim \sigma \cdot T^{\frac{1}{2}} \quad (11)$$

$$e^{-\frac{e}{k T}} \sim [\sigma \cdot T^{\frac{1}{2}}] \bar{U}_i. \quad (11a)$$

Für die Emission der Linienstrahlung ergibt sich aus der Boltzmann-Verteilung die Beziehung:

$$S_L \sim p T^{-1} e^{-\frac{e U_A}{k T}}. \quad (13)$$

Also wird mit Gl. (11a):

$$\frac{U_A}{U_i} = \frac{1}{2} \frac{\ln \frac{S_L}{S_{L1}} + \ln \frac{T_1}{T_2}}{\ln \sigma_2 - \ln \sigma_1}. \quad (14)$$

Gl. (14) gibt uns die Möglichkeit, die effektive Ionisierungsspannung durch Messungen der Intensität bestimmter Linien des Spektrums zu bestimmen, bei denen U_A genau bekannt ist.

Einfluß der Reabsorption auf die Anregungsspannung der Linien

Hierbei muß jedoch der Einfluß der Reabsorption und Selbstumkehr der Linien beachtet werden. Wir wollen auf den Einfluß der Reabsorption näher eingehen. Es läßt sich zeigen, daß für eine durch Reabsorption beeinflusste Linienstrahlung nicht die Beziehung der Gl. (13) gilt, sondern daß $e^{-\frac{e U_A}{k T}}$ übergeht in $e^{-\frac{e U_{(K \cdot l)}}{k T}}$, wobei $U_{(K \cdot l)} < U_A$ ist. $K \cdot l$ ist das Produkt aus Absorptionskoeffizienten K und Schichtdicke l , innerhalb deren die Reabsorption stattfindet. Im allgemeinen

wird die Absorption bzw. Emission eines Atoms durch benachbarte Atome oder Moleküle gestört. Wir wollen jedoch unsere Betrachtung auf ungestörte Atome beschränken.

Für die in der Spektrallinie absorbierte Energie pro sec und cm^3 erhält man^{a)}

$$\frac{\varepsilon_\nu}{\int k_\nu d\nu} = c \cdot U_\nu = 4\pi J_\nu. \quad (15)$$

Das ist eine der Kirchhoffschen ähnliche Beziehung, die hier nur noch für die gesamte Strahlung einer Spektrallinie gilt und nicht, wie im Fall der schwarzen Strahlung, für jedes beliebige Frequenzgebiet. Dies ist auch verständlich, denn ein Strahlungsgleichgewicht ist mit der Existenz von Spektrallinien unverträglich.

Wir erhalten die gesamte Emission von einer Spektrallinie der Frequenz ν aus einer Schicht der Länge l , indem wir die in der gleichen Schicht stattfindende Absorption berechnen. Diese ist

$$\int_0^l J_\nu \cdot K_\nu \cdot e^{-K_\nu \cdot l} dl = J_\nu (1 - e^{-K_\nu \cdot l}). \quad (16)$$

Für $K_\nu \cdot l \ll 1$ wird $1 - e^{-K_\nu \cdot l} \rightarrow K_\nu \cdot l$ und wir erhalten auf die Volumeneinheit bezogen

$$C \cdot U_\nu \cdot K_\nu \cdot l = \varepsilon_\nu \cdot l. \quad (17)$$

Unter ε_ν ist die pro sec erfolgende Ausstrahlung der Volumeneinheit verstanden. Unter Vernachlässigung der erzwungenen Übergänge erhalten wir nach Einstein^{b)} für die spontane Emission

$$\varepsilon_\nu = h \cdot \nu_{nm} \cdot A_{nm} \cdot N_m, \quad (18)$$

wobei A_{nm} der Koeffizient für die Übergangswahrscheinlichkeit $m \rightarrow n$ und

$$N_m = \frac{g_m}{g_0} \cdot N_0 \cdot e^{-\frac{E_m}{kT}} \quad (14)$$

die Anzahl der Atome im oberen Zustand ist. m und n sind Symbole für sämtliche Quantenzahlen.

Da $E_m = e U_A$ ist, ist

$$\varepsilon_\nu = \frac{g_m}{g_0} \cdot N_0 \cdot h \nu \cdot e^{-\frac{e U_A}{kT}}. \quad (19)$$

Die gesamte Strahlung einer Spektrallinie der Frequenz ν aus einer Schicht der Länge l wird für den Fall

1. $K_\nu \cdot l \ll 1$

$$S_\nu = \text{const} \cdot e^{-\frac{e U_A}{kT}}, \quad (20)$$

^{a)} Vgl. A. Unsöld, Physik der Sternatmosphären IX. 45, Berlin 1938.

^{b)} A. Einstein, Physik. Z. 18, 121 (1917).

für den Fall

2. $K_p \cdot l \gg 1$ wird $(1 - e^{-K_p \cdot l}) \rightarrow 1$, also

$$S_p = e \cdot U_p = \text{const} \cdot e^{-\frac{h\nu}{kT}} \quad (21)$$

gleich der Strahlung des schwarzen Körpers entsprechend dem Wienschen Strahlungsgesetz.

Definieren wir jetzt ein $\mathfrak{E}_{(K_p, n)} = e U_{(K_p, n)}$ durch den für beliebige $K_p \cdot l$ geltenden Ansatz:

$$e^{-\frac{h\nu}{kT}} (1 - e^{-K_p \cdot l}) = e^{-\frac{e[U_{(K_p, n)}]}{kT}},$$

so ist

$$h\nu < e U_{(K_p, n)} < e U_A.$$

Der sich ergebende Wert für die „Anregungsspannung“ wird, wie hieraus deutlich wird, um so kleiner, je größer die Reabsorption ist.

Die Größe der Reabsorption hängt von mehreren Faktoren ab. Vor allem ist sie abhängig von den Besetzungszahlen der Ausgangsniveaus. Wir müssen also für unsere Messungen die auf den niedrigen 2^3P_2 , 2^3P_1 - und 2^3P_0 -Niveaus endenden Linien, die durch Selbstumkehr gestört sind, von vornherein ausscheiden. Aber auch bei auf oberen Niveaus endenden Linien wird der Einfluß der Reabsorption nicht ohne weiteres zu vernachlässigen sein, da bei dem großen Druck von etwa 60 atm ($5 \cdot 10^4$ Torr.) die Besetzungszahlen auch der höheren Terme große Werte erreichen¹⁰⁾.

Das Absorptionsvermögen ist abhängig von der Intensitätsverteilung, also von der Art der Verbreiterung der Linie¹¹⁾. Die Eigendruckverbreiterung ist unabhängig von der Leitfähigkeit, jedoch ergibt sich für die Elektronen- bzw. Ionenstoßverbreiterung eine Stromstärkeabhängigkeit.

Messungen an Linien ohne Reabsorption

Will man das Verhältnis der Anregungsspannung zur Ionisierungsspannung für die Linien auf Grund von Intensitäts- und Leitfähigkeitsmessungen bestimmen, so können hierfür im wesentlichen nur die ultraroten Linien, insbesondere die Linie 10140 Å, herangezogen werden.

Die Messungen wurden für die Linien 10140 Å, 11287 Å, 12070 Å und 13673 Å durchgeführt. Die Trennung der Linien erfolgte mittels Doppelmonochromator von Halle Nachfolger, Berlin-Steglitz, mit Quarzprismen. Der Einfluß des Kontinuums, der bei der geringen Dispersion des Doppelmonochromators in diesem Gebiet sehr groß ist, wurde nach der von Rössler¹²⁾ angegebenen Methode (Messungen bei verschiedenen Spaltbreiten des Monochromators) bestimmt und entsprechend eliminiert. Es wurde mit einem Eintrittsspalt von 0,15 mm und einem Mittelspalt von 0,3 mm gearbeitet. Zur Messung wurde ein Vakuum-Thermoelement von Müller mit 0,5 mm Breite benutzt. Die Thermostrome wurden mit

¹⁰⁾ K. H. Riewe u. R. Rompe, Z. Physik **111**, 79 (1938).

¹¹⁾ R. Ladenburg u. F. Reiche, Z. Physik **65**, 200 (1930).

¹²⁾ F. Rössler, Z. Physik **110**, 495 (1938).

einem schnell schwingenden Spiegelgalvanometer von Ruhstrat-Göttingen — Stromempfindlichkeit $5 \cdot 10^{-8}$ Amp. — in willkürlichen Einheiten bestimmt.

Die Messungen wurden an Quecksilberhochdrucklampen der Typen HBO 500 und HBO 200¹³⁾ durchgeführt. Der Druck in der Lampe betrug etwa 35 atm bei der HBO 500 und etwa 60 atm bei der HBO 200. Die Messungen erfolgten, indem die bei Sollast brennenden Lampen so kurzzeitig unter- bzw. überlastet wurden,

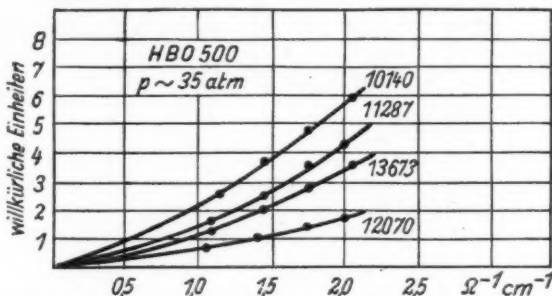


Abb. 3. Strahlungsintensität der vier ultravioletten Linien in Abhängigkeit von der Leitfähigkeit bei einem Druck von $p \sim 35$ atm.

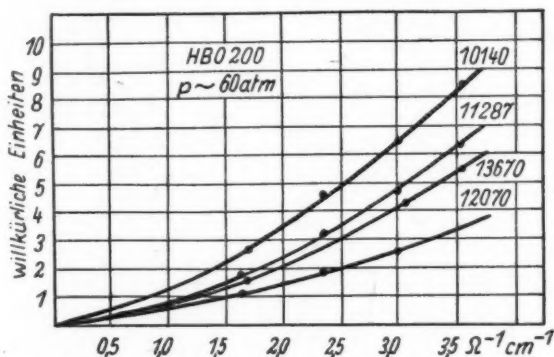


Abb. 4. Strahlungsintensität der vier ultravioletten Linien in Abhängigkeit von der Leitfähigkeit bei einem Druck von $p \sim 60$ atm.

daß der Druck während der Meßreihe als konstant angenommen werden konnte. Dies zeigt auch die der Charakteristik entsprechende geringe Zunahme der Feldstärke. Sämtliche Messungen wurden fünfmal durchgeführt; die Einzelwerte zeigten sehr gute Übereinstimmung. Die elektrischen Daten sind aus Tabelle 1 zu ersehen. In den Abb. 3 und 4 sind die gemessenen Intensitäten in Abhängigkeit von der Leitfähigkeit aufgetragen. Die sich hieraus errechnenden Werte für das Verhältnis $\frac{U_A}{U_i}$ und die sich daraus ergebenden Werte für U_i sind in Tabelle 2 aufgeführt.

¹³⁾ R. Rompe u. W. Thouret, Das Licht, 7—10, 1944.

Tabelle 2

Werte für $\frac{U_A}{U_i}$ für die ultraroten Hg-Linien
 Werte für die effektive Ionisierungsspannung
 I HBO 500 p ~ 35 atm

Wellenlänge Å	Anregungs- spannung Volt	Leistungskonzentration kW/cm ³			
		17—22	22—29	29—38	38—51
		Leitfähigkeit $\Omega^{-1} \text{ cm}^{-1}$			
		1,0—1,2	1,2—1,45	1,45—1,75	1,75—2,1
Werte für $\frac{U_A}{U_i}$					
10140	7,9	0,77	0,77	0,76	0,76
11287	8,8	0,84	0,85	0,87	0,90
12070	7,7	0,76	0,76	0,79	0,76
13673	8,65	0,86	0,85	0,85	0,88
Werte für U_i					
10140	7,9	10,25	10,25	10,4	10,4
11287	8,8	(10,5)	10,3	10,1	9,8
12070	7,7	10,1	10,1	9,8	10,1
13673	8,65	10,1	10,2	10,2	9,8

II HBO 200 p ~ 60 atm

Wellenlänge Å	Anregungs- spannung Volt	Leistungskonzentration kW/cm ³				
		43-54	54-65	65-85	85-106	106-140
		Leitfähigkeit Ω ⁻¹ cm ⁻¹				
		1,5-1,8	1,8-2,15	2,15-2,6	2,6-3,1	3,1-3,7
Werte für $\frac{U_A}{U_i}$						
10140	7,9	0,77	0,78	0,78	0,78	0,79
11287	8,8	0,86	0,86	0,86	0,87	0,89
12070	7,7	0,78	0,78	0,78	0,78	0,8
13673	8,65	0,84	0,84	0,85	0,85	0,86
Werte für U_i						
10140	7,9	10,25	10,1	10,1	10,1	10,0
11287	8,8	10,25	10,25	10,25	10,1	9,9
12070	7,7	9,9	9,9	9,9	9,9	9,6
13673	8,65	10,3	10,3	10,2	10,2	10,1

Werte für die effektive Ionisierungsspannung

Die Streuung der Einzelwerte für die effektive Ionisierungsspannung ist sehr groß, jedoch liegen die gefundenen Werte bis auf einen Wert bei der Linie 12070 bei 60 atm oberhalb 9,7 Volt. Die Meßergebnisse dieser Linie bei 60 atm liegen allgemein an der unteren Streugrenze. Wegen der geringen Intensität dieser Linie ist die Fehlermöglichkeit größer. Trägt man die Ergebnisse beider Meßreihen in Abhängigkeit von der Leistungskonzentration auf, wie es Abb. 5 zeigt, so erkennt man eine Abnahme der Ionisierungsspannung mit wachsender Leistung. Die Herabsetzung der Ionisierungsgrenze ist im allgemeinen gering, sie ist bestimmt kleiner als 1 Volt, wie sie auf Grund der Unsöld'schen Gleichung für die kontinuierliche Strahlung angenommen wurde. Übereinstimmend mit den eingangs angeführten theoretischen Überlegungen erhalten wir für die Ionisierungsspannung

in der Hochdruckentladung für Leistungskonzentrationen $< 50 \text{ kW/cm}^2$ einen Wert von 10,2 Volt. Dieser Wert nimmt mit wachsender Leistung (Temperatur) weiterhin ab und erreicht bei einer Leistungskonzentration von 150 kW/cm^2 etwa den Wert 9,9 Volt.

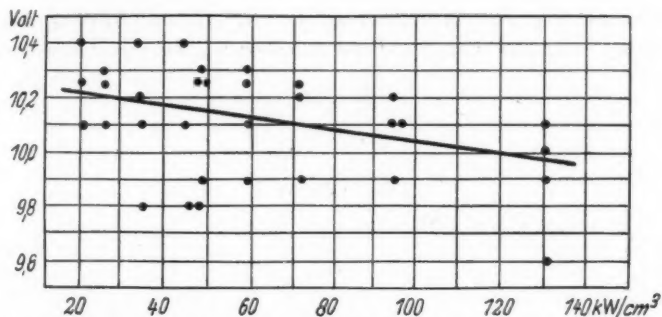


Abb. 5. Effektive Ionisierungsspannung der Hg-Hochdruckentladung in Abhängigkeit von der Leistungskonzentration

Berlin-Buch, Institut für Festkörperforschung der Deutschen Akademie der Wissenschaften zu Berlin.

(Bei der Redaktion eingegangen am 1. Februar 1950.)

Die Entstehung des Kontinuums in der Quecksilberhochdruckentladung und die Größe ihrer mittleren „Anregungsspannung“

Von R. Borchert

(Mit 1 Abbildung)

Inhaltsübersicht

Zusammenfassend läßt sich sagen, daß die kontinuierliche Strahlung in der Quecksilberhochdruckentladung im allgemeinen etwa zu gleichen Teilen aus Stoßmolekül- und Rekombinationsstrahlung entsteht. Trotz der quadratischen Abhängigkeit der Stoßmolekülstrahlung vom Druck wird bei hohen Drucken jedoch bei entsprechend großen Leistungen auch ein Zustand erreicht, in dem die Rekombinationsstrahlung den überwiegenden Anteil am Kontinuum aufweist. Das erklärt sich aus der Zunahme der Elektronenkonzentration, die auf zwei Effekten beruht: auf der Erhöhung der Temperatur und auf dem Freiwerden von Elektronen innerhalb der oberen Terme. Die durch den letzten Effekt bewirkte Herabsetzung der Ionisierungsspannung ist zwar gering, jedoch ist in dem vorliegenden Temperaturbereich (8000–9000°) dieser zusätzliche Anteil an Rekombinationsstrahlung bereits bei einer Herabsetzung von $U = 0,4$ – $0,5$ Volt von der gleichen Größe, wie die Intensität von Rekombinations- und Bremsstrahlung, wie sie sich für die Ionisierungsspannung des isolierten Atoms ergibt. Dies errechnet sich aus dem von Unsöld angegebenen Ausdruck für das Verhältnis beider Anteile:

$$\frac{\text{Anteil kontinuierlich-kontinuierlich} + \text{kontinuierlich-diskret}}{\text{Anteil der verschmierten Terme}} = \frac{1}{\frac{\Delta U}{e^{kT}} - 1}.$$

Einleitung: Entstehung des Kontinuums

Das Kontinuum in der Quecksilberhochdruckentladung kann nicht auf die Strahlung stabiler Moleküle zurückgeführt werden, da wegen der geringen Dissoziationsenergie bei den vorliegenden Temperaturen alle Moleküle dissoziiert sind. Kontinuierliche Strahlung kann hier entstehen durch Strahlung von Stoßmolekülen¹⁾ und durch Rekombinations- und Bremsstrahlung²⁾ (im folgenden kurz Rekombinationsstrahlung genannt, da die Intensität der Bremsstrahlung wesentlich kleiner ist).

Energetische Abschätzungen ergeben für beide Anteile keine Übereinstimmung mit den gemessenen Intensitäten des Kontinuums. Hier weist Unsöld³⁾ darauf hin, daß die Übergänge aus den dicht unterhalb der Ionisierungsgrenze liegenden ver-

¹⁾ F. Rössler, Z. Physik **112**, 667 (1939); **122**, 285 (1944).

²⁾ W. Finkelburg, Kontinuierliche Spektren, Berlin 1938.

³⁾ A. Unsöld, Ann. Physik **33**, 607 (1938).

schmierten Terme ebenfalls kontinuierliche Emission ergeben, die er als Pseudo-Kontinuum bezeichnet. Bei der Aufstellung der quantitativen Formel für die Emission bezieht er diese Übergänge in die Rekombinationsstrahlung mit ein:

$$\varepsilon_p = \text{const} \cdot p \cdot T \cdot e^{-\frac{e(U_i - \Delta U)}{kT}} \quad (1)$$

ΔU ist der Bereich der verschmierten Terme. Da diese Gleichung gute Übereinstimmung mit gemessenen Intensitätswerten ergibt, wenn man für ΔU etwa 1 Volt annimmt, so wurde gefolgert, daß das Kontinuum in der Hochdruckentladung im wesentlichen auf Rekombinationsstrahlung beruhe⁴⁾. Das bedingt aber eine Herabsetzung der Ionisierungsgrenze um etwa 1 Volt, wie sie tatsächlich nicht vorliegt⁵⁾.

Da die Intensität der Stoßmolekül-Strahlung entsprechend der Beziehung

$$S_M = \text{const} \cdot p^2 \cdot T^{-1,5} \cdot e^{-\frac{eU_M}{kT}} \quad (2)$$

quadratisch mit dem Druck ansteigt, kommt Elenbaas⁶⁾ auf Grund von Messungen bei wandstabilisierten Bogen geringer Leistung (35 W/cm) und konstanter Temperatur zu der Auffassung, daß bei hohen Drucken der Anteil der Stoßmolekül-Strahlung größer als der der Rekombinationsstrahlung ist. Experimentell findet man diese starke Druckabhängigkeit nicht. Nach den von Elenbaas angegebenen Werten verhält sich der Anteil der Stoßmolekül-Strahlung zur Rekombinationsstrahlung bei einer Bogenleistung von 35 W/cm im Bereich bis zu 6 atm wie 2,5:1, im Bereich 8–10 atm wie 1,5:1.

Problemstellung

Da die Anregungsspannung U_M für die Stoßmolekül-Strahlung wesentlich kleiner ist als die Ionisierungsspannung, geht bei Leistungserhöhung die Temperatur in die Gleichung für die Intensität bei der Rekombinationsstrahlung stärker ein als bei Stoßmolekül-Strahlung. Hinzu kommt noch die Verminderung der effektiven Ionisierungsspannung, so daß der Anteil der Rekombinationsstrahlung bei Bogen hoher Leistung und hoher Temperatur größer wird. Das Verhältnis von Stoßmolekül-Strahlung zu Rekombinationsstrahlung ist

$$\frac{S_M}{S_R} = \frac{C_M}{C_R} \cdot p \cdot T^{-2,5} \cdot e^{-\frac{e(U_i - U_M)}{kT}},$$

es ist also neben dem Druck abhängig von den Exponenten der e -Funktion, in Sonderheit von der Temperatur.

Rompe und Steenbeck⁷⁾ haben den Betrag des Molekülkontinuums berechnet und nur einen Wert von etwa 10^{-3} der experimentell gemessenen Intensität erhalten. Demnach wäre der Anteil der Stoßmolekül-Strahlung allgemein vernachlässigbar gering. Andererseits ist die Herabsetzung der Ionisierungsspannung nicht so groß, um die Entstehung des Kontinuums in der Hochdruckentladung durch Rekombination deuten zu können.

Um zur Klärung dieses Problems beizutragen, werden Messungen an Bögen extrem hoher Leistungskonzentrationen bei hohen Drucken durchgeführt.

⁴⁾ P. Schulz, Z. Physik **119**, 167 (1942).

⁵⁾ R. Borchert, Ann. Physik (6) **7**, 3.1 (1950).

⁶⁾ W. Elenbaas, Philips Res. Rep. **2**, 442 (1947).

⁷⁾ R. Rompe u. M. Steenbeck, Ergebn. exakt. Naturw. **18**, 257 (1939).

Strahldichte als Funktion von Leitfähigkeit und dem Verhältnis Anregungs- zu Ionisierungsspannung

Für konstanten Druck läßt sich wegen der geringen Temperaturänderung für die Intensität des Gesamtkontinuums die Beziehung ansetzen:

$$S_k \sim e^{-\frac{eU_M}{kT}}, \quad (3)$$

wobei U_k einen statistischen Mittelwert der Anregungsspannung entsprechend den verschiedenen Anteilen darstellt.

Da die Elektronenkonzentration n_e proportional der Leitfähigkeit σ ist und n_e sich aus der Saha-Gleichung bestimmen läßt, sofern für U_i die effektive Ionisierungsspannung eingesetzt wird, ergibt sich für den Ausdruck

$$e^{-\frac{e}{kT}} \sim \sigma^{\frac{2}{U_i}}. \quad (4)$$

Aus Gl. (3) und (4) erhalten wir eine Beziehung für das Verhältnis von $\frac{U_k}{U_i}$:

$$S_k = \text{const} \cdot \sigma^{\frac{2U_k}{U_i}} \quad (5)$$

$$\frac{U_k}{U_i} = \frac{1}{2} \frac{\ln S_{k2} - \ln S_{k1}}{\ln \sigma_2 - \ln \sigma_1}. \quad (5a)$$

Intensität S_k und Leitfähigkeit $\sigma = \frac{J}{\pi R^2 \mathcal{E}}$ können experimentell bestimmt werden.

Experimentelle Durchführung der Messungen

Die Messungen wurden an Quecksilberhöchstdrucklampen HBO 500 und HBO 200^{a)} bei Drucken von 35 atm bzw. 60 atm durchgeführt. Für die Elektronenverluste wurden auf Grund von früheren Messungen mit Lampen von Spannungen von 10–15 Volt (Type HBO 107) 10 Volt angenommen. Die Änderung der Leistung erfolgt nur für den kurzen Augenblick der Messung (20 sec), so daß eine Druckänderung durch Kondensation oder Verdampfen von Quecksilber nicht eintreten konnte. Eine Vergleichsmessung ergab, daß eine Erhöhung des Druckes von mehr als 1% bei Erhöhung der zugeführten Leistung erst nach > 1 min auftritt, da erst nach dieser Zeit eine entsprechende Aufheizung der Kolbenträume, in denen sich nicht verdampfendes Quecksilber befinden kann, erfolgt ist. Die gemessene Erhöhung der Feldstärke mit wachsender Leistung entspricht der Beziehung $\mathcal{E} \sim J^{\frac{1}{2}} - J^{\frac{1}{2}}$. Elenbaas hat eine Beziehung $\mathcal{E} \sim J^{\frac{1}{2}}$ für niedrigere Drucke gefunden. Jedenfalls entspricht die Änderung der Feldstärke der normalen Charakteristik und läßt darauf schließen, daß der Druck während der Messungen konstant geblieben ist.

Für die Messung der Strahlungsintensität wurde ein Doppelmonochromator von Halle-Nachfolger, Berlin-Steglitz, mit Quarzprismen benutzt. Der Eintrittsspalt wurde auf $0,25 \times 10$ mm eingestellt, der Mittelspalt auf 0,5 mm. In die Austrittsöffnung wurde ein Vakuum-Thermoelement von Müller mit einer Breite von 0,5 mm gesetzt. Die Thermostrome wurden in willkürlichen Einheiten mittels eines Schnellschwingers von Ruhstrat-Göttingen, Stromempfindlichkeit $5 \cdot 10^{-9}$ Amp. gemessen. Die Intensität des Kontinuums wurde an zwölf Stellen

^{a)} R. Rompe u. W. Thouret, Das Licht, 7–10, 1944.

im Bereich von 3370–8150 Å für vier verschiedene Leitfähigkeiten bestimmt. Die Wellenlängen wurden so gewählt, daß ein Einfluß der Linien möglichst ausgeschaltet wurde. Sämtliche Messungen wurden dreimal durchgeführt. Die Übereinstimmung war außerordentlich gut. Die Mittelwerte sind in Tabelle 1 und 2 zusammengestellt. Die Meßwerte wurden im Falle der HBO 200 an fünf Stellen des Kontinuums bei kurzer Spaltlänge (1,0 mm) kontrolliert, um zu ersehen, ob die Brennflecke an den Elektroden spitzen die Meßwerte beeinflusst haben. Es zeigte sich keine systematische Abweichung, da die Spaltlänge von 10 mm bei der vergrößert abgebildeten Lichtbogensäule (Elektrodenabstand 2,5 mm) diesen Einfluß schon ausschließt.

Meßergebnisse. Werte für U_k

Die Tabellen 1 und 2 enthalten die gemessenen Werte für Leitfähigkeit und Strahlungsintensität. Die Auswertung dieser Meßwerte geschah auf folgende Weise: Die Intensität wurde für jede Wellenlänge als Funktion der Leitfähigkeit aufgetragen. Aus den so gezeichneten Kurven wurde die Änderung der Intensität für konstante Leitfähigkeitsbereiche $\sigma_2:\sigma_1 = 1,2$ abgelesen und daraus der Exponent $\frac{U_k}{U_i}$ berechnet. Die so bestimmten Werte $\frac{U_k}{U_i}$ sind in den Tabellen 3 und 4 aufgeführt. Sie zeigen bei einer relativ großen Streuung von $\pm 0,1$ keine Frequenzabhängigkeit. Das entspricht auch dem Unsöld'schen Ansatz (Gl. 1) für die Rekombinationsstrahlung, während für Stoßmolekülstrahlung frequenzunabhängige Werte nicht zu erwarten sind.

Wie in Tabelle 4 aufgeführt, sind bei Leistungskonzentrationen $> 100 \text{ kW/cm}^2$ Werte für $\frac{U_k}{U_i} > 1$ gefunden worden. Das läßt sich aus der Vernachlässigung der Temperaturzunahme erklären. Gl. (5a) müßte für Rekombinationsstrahlung genauer heißen:

$$\frac{U_k}{U_i} = \frac{1}{2} \frac{\ln \frac{S_2}{S_1} - \ln \frac{T_2}{T_1}}{\ln \sigma_2 - \ln \sigma_1} \quad (5b)$$

Bei einem Druck von 35 atm und Leistungskonzentrationen von 20–50 kW/cm² erhalten wir Werte für das Verhältnis U_k und U_i , die unter 1 liegen. Bei Drucken von 60 atm erreicht man bei einer Leistungskonzentration von 100 kW/cm² gleiche Werte für U_k und U_i . Erst für diesen extremen Fall kann man annehmen, daß das Kontinuum praktisch überwiegend durch Rekombinationsstrahlung entsteht. Im allgemeinen ist jedoch der Anteil der Stoßmolekülstrahlung wesentlich größer, als auf Grund der Unsöld'schen Formel angenommen wurde.

Rössler¹⁾ hat für die Kapillar-Höchstdrucklampe HgF 700^o), die eine Leistungskonzentration von $\sim 50 \text{ kW/cm}^2$ und eine Leitfähigkeit von $\sim 0,8 \Omega^{-1} \text{ cm}^{-1}$ bei einem Druck von 100 atm aufweist, einen Wert von 8,4–8,7 Volt gefunden. Stellen wir die Werte für annähernd gleiche Leistung bei den verschiedenen Drucken gegenüber, so ergibt sich für:

$$\begin{aligned} L = 50 \text{ kW/cm}^2: p \sim 35 \text{ atm}: U_k &= 9,8 \text{ Volt,} \\ p \sim 60 \text{ atm}: U_k &= 9,2 \text{ Volt,} \\ p \sim 100 \text{ atm}: U_k &= 8,6 \text{ Volt.} \end{aligned}$$

¹⁾ J. Kern, Z. techn. Physik **20**, 250 (1939); **23**, 108 (1942).

Tabelle 1

Meßwerte der elektrischen Daten und der Strahlungsdichte des Kontinuums einer Quecksilberhöchstdrucklampe HBO 500 bei verschiedener Stromstärke. Betriebsdruck 35 atm = const. Monochromator-Eintrittsspalt $0,25 \times 10,0$ mm

λ	J	R_0	$\frac{J}{\pi R_0^2}$	U	\mathcal{E}	L	σ	α
Å	Amp	10^{-2} cm	Amp/cm ²	Volt	Volt/cm	kW/cm ²	Ω^{-1} cm ⁻¹	SKT
3370	3,5	9,1	134,5	70,5	130	17,5	1,035	0,7
	5	9,2	188	74	137,5	25,85	1,375	1,2
	6,5	9,2	244	76,5	143	34,9	1,71	1,7
	8	9,3	294	79	148,5	43,6	1,98	2,3
3420	3,5	9,1	134,5	70	129	17,35	1,04	0,5
	5	9,2	188	73,5	136,5	25,65	1,38	0,9
	6,5	9,2	244	76,5	143	34,9	1,71	1,3
	8	9,3	294	78,5	147,5	43,3	2,0	1,8
3550	3,5	9,1	134,5	71,5	132,3	17,8	1,015	1,0
	5	9,2	188	73,5	136,5	25,65	1,38	1,7
	6,5	9,2	244	76,5	143	34,9	1,71	2,2
	8	9,3	294	79	148,5	43,6	1,98	3,3
3750	3,5	9,1	134,5	70,5	130	17,5	1,035	1,2
	5	9,2	188	73,5	136,5	25,65	1,38	2,3
	6,5	9,2	244	76	142	34,65	1,72	3,5
	8	9,3	294	78,5	147,5	43,3	2,0	4,9
3900	3,5	9,1	134,5	70	129	17,35	1,04	1,0
	5	9,2	188	73,5	136,5	25,65	1,38	1,8
	6,5	9,2	244	76,5	143	34,9	1,71	2,6
	8	9,3	294	79	148,5	43,6	1,99	4
3980	3,5	9,1	134,5	69,5	127,9	17,2	1,05	0,7
	5	9,2	188	73,5	136,5	25,65	1,38	1,1
	6,5	9,2	244	75,5	140,8	34,35	1,73	1,6
	8	9,3	294	78	146	43,0	2,01	2,2
4500	3,5	9,1	134,5	70	129	17,35	1,04	1,0
	5	9,2	188	73	135,5	25,5	1,39	1,6
	6,5	9,2	244	76	141,9	34,65	1,72	2,4
	8	9,3	294	78	146	43,0	2,01	3,3
4770	3,5	9,1	134,5	70	129	17,35	1,04	0,9
	5	9,2	188	73,5	136,5	25,65	1,38	1,6
	6,5	9,2	244	76	141,9	34,65	1,72	2,4
	8	9,3	294	78,5	147,3	43,3	2,0	3,4
5220	3,5	9,1	134,5	69	126,9	17,05	1,06	0,8
	5	9,2	188	73,6	136,7	25,7	1,38	1,4
	6,5	9,2	244	75,5	140,8	34,35	1,73	2,0
	8	9,3	294	78	146,2	43,0	2,01	2,7
6370	3,5	9,1	134,5	69,5	127,9	17,2	1,05	1,1
	5	9,2	188	73,5	136,5	25,65	1,38	2,1
	6,5	9,2	244	75,5	140,8	34,35	1,73	2,9
	8	9,3	294	78,5	147,3	43,3	2,0	4,0
6970	3,5	9,1	134,5	69,5	127,9	17,2	1,05	1,3
	5	9,2	188	73,5	136,5	25,65	1,38	2,3
	6,5	9,2	244	76	141,9	34,65	1,72	3,2
	8	9,3	294	78	146,2	43,0	2,01	4,3
8150	3,5	9,1	134,5	69,5	127,9	17,2	1,05	0,9
	5	9,2	188	73,5	136,5	25,65	1,38	1,6
	6,5	9,2	244	76	141,9	34,65	1,72	2,2
	8	9,3	294	78	146,2	43,0	2,01	3,0

nuums
Strom-
esspalt

α

SKT

0,7
1,2
1,7
2,3
0,5
0,9
1,3
1,8
1,0
1,7
2,2
3,3
1,2
2,3
3,5
4,9
1,0
1,8
2,6
1
0,7
1,1
1,6
2,2
1,0
1,6
1,4
1,3
1,9
1,6
1,4
1,4
1,8
1,4
1,0
1,7
1,1
1,1
1,9
1,0
1,3
1,3
2
1,3
1,9
1,6
2
2
0

M.
eil
at

33

34

35

375

390

398

4500

4700

5220

6380

6970

8150

Ann.

Tabelle 2a

Meßwerte der elektrischen Daten und der Strahlungsdichte des Kontinuums einer Quecksilberhöchstdrucklampe HBO 200 bei verschiedener Stromstärke, Betriebsdruck 60 atm = const. Monochromator-Eintrittsspalt $0,25 \times 10,0$ mm

λ	J	R_0	$\frac{J}{\pi R_0^2}$	U	\mathcal{E}	L	σ	α
Å	Amp	10^{-2} cm	Amp/cm ²	Volt	Volt/cm	kW/cm ³	Ω^{-1} cm ⁻¹	SKT
3370	2	$4,3 \cdot 10^{-2}$	345	53,1	172,5	59,4	2,0	1,6
	3	4,4	492	56,2	185	91,0	2,66	2,7
	4	4,5	630	57,8	191	120,4	3,3	4,4
	5	4,6	757	59,3	197	149	3,84	6,5
3440	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	0,9
	3	4,4	492	55,8	183	90,0	2,68	1,4
	4	4,5	630	57,5	190	119,6	3,32	2,4
	5	4,6	757	59,7	199	150,5	3,80	3,6
3570	2	$4,3 \cdot 10^{-2}$	345	53,5	174	60,0	1,98	1,4
	3	4,4	492	55,2	181	89,1	2,72	2,3
	4	4,5	630	57,1	188,5	118,7	3,34	3,9
	5	4,6	757	59,4	197,5	149,5	3,83	5,5
3750	2	$4,3 \cdot 10^{-2}$	345	53,2	172,7	59,6	2,0	2,7
	3	4,4	492	56,0	184	90,5	2,67	5,0
	4	4,5	630	57,4	189,5	119,4	3,32	8,0
	5	4,6	757	59,5	198	150	3,82	11,6
3900	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	2,0
	3	4,4	492	56,1	185	90,8	2,66	3,8
	4	4,5	630	57,5	189	119,6	3,32	6,0
	5	4,6	757	59,5	198	150	3,82	8,4
3980	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	1,8
	3	4,4	492	56,2	185	91,0	2,66	3,2
	4	4,5	630	57,3	189	119	3,33	5,4
	5	4,6	757	59,5	198	150	3,82	7,6
4500	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	2,4
	3	4,4	492	56,2	185	91,0	2,66	4,7
	4	4,5	630	57,3	189	119	3,33	7,4
	5	4,6	757	59,5	198	150	3,82	10,7
4700	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	2,3
	3	4,4	492	55,8	183	90,0	2,68	4,2
	4	4,5	630	57,4	189,5	119,4	3,32	6,7
	5	4,6	757	59,5	198	150	3,82	9,4
5220	2	$4,3 \cdot 10^{-2}$	345	53,1	172,3	59,4	2,0	2,3
	3	4,4	492	55,8	183	90,0	2,68	3,8
	4	4,5	630	57,1	188,3	118,5	3,35	5,6
	5	4,6	757	59,5	198	150	3,82	8,2
6380	2	$4,3 \cdot 10^{-2}$	345	53,1	172,3	59,4	2,0	3,8
	3	4,4	492	55,7	182,7	90,0	2,68	6,5
	4	4,5	630	57,1	188,3	118,3	3,35	10,0
	5	4,6	757	59,5	198	150	3,82	14,4
6970	2	$4,3 \cdot 10^{-2}$	345	53,4	173,5	59,8	1,99	3,8
	3	4,4	492	55,6	182,4	89,7	2,7	6,9
	4	4,5	630	57,2	188,5	118,7	3,34	10,7
	5	4,6	757	59,7	198,6	150,5	3,81	15,1
8150	2	$4,3 \cdot 10^{-2}$	345	53,2	172,6	59,6	2,0	2,8
	3	4,4	492	55,8	183	90,0	2,68	4,9
	4	4,5	630	57,2	188,5	118,7	3,34	7,5
	5	4,6	757	59,5	198	150	3,82	11,6

Tabelle 2b

Meßwerte der elektrischen Daten und der Strahlungsdichte des Kontinuums einer Quecksilberhöchstdrucklampe HBO 200 bei verschiedener Stromstärke. Betriebsdruck 60 atm = const. Monochromator-Eintrittsspalt $0,25 \times 1,0$ mm

λ	J	R_0	$\frac{J}{\pi R_0^2}$	U	\mathcal{E}	L	σ	α
Å	Amp.	10^{-2} cm	Amp/cm ²	Volt	Volt/cm	kW/cm ²	Ω^{-1} cm ⁻¹	SKT
3750	2	$4,3 \cdot 10^{-2}$	345	52,5	170	58,7	2,06	1,2
	3	4,4	492	54,7	179	88,1	2,75	2,0
	4	4,5	630	56,9	187,5	118	3,36	3,4
	5	4,6	757	59,2	197	149	3,84	5,0
3900	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	0,8
	3	4,4	492	55,6	182,4	89,7	2,7	1,6
	4	4,5	630	57,1	188,3	118,5	3,35	2,3
	5	4,6	757	59,3	197	149	3,84	3,4
4500	2	$4,3 \cdot 10^{-2}$	345	53,3	173	59,7	2,0	1,1
	3	4,4	492	55,8	183	90,3	2,69	1,9
	4	4,5	630	57,3	189	119	3,33	3,1
	5	4,6	757	59,3	197	149	3,84	4,4
4700	2	$4,3 \cdot 10^{-2}$	345	52,3	169	58,3	2,08	0,9
	3	4,4	492	54,8	179	88	2,74	1,6
	4	4,5	630	56,9	187,5	118	3,36	2,6
	5	4,6	757	59,1	196,5	148,7	3,85	3,7
6380	2	$4,3 \cdot 10^{-2}$	345	53,2	172,7	59,6	2,0	1,2
	3	4,4	492	55,8	183	90,3	2,69	2,1
	4	4,5	630	57,1	188,5	118,8	3,34	3,3
	5	4,6	757	59,2	197	149	3,84	4,7

Tabelle 3

Werte für $\frac{U_k}{U_i} \cdot p \sim 35$ atm = const

Wellenlänge λ Å	Leistungskonzentration kW/cm ²			
	17-22	22-29	29-38	38-51
	Leitfähigkeit Ω^{-1} cm ⁻¹			
	1,0-1,2	1,2-1,44	1,44-1,75	1,75-2,1
3370	0,90	0,90	0,92	0,92
3420	0,96	0,92	0,94	0,94
3550	0,76	0,77	0,94	> 1 starke Streuung
3750	1,0	0,98	1,0	
3900	0,78	0,80	0,96	1,01
3980	0,88	0,82	0,92	1,0
4500	0,85	0,78	0,96	0,96
4770	0,96	0,94	0,96	0,96
5220	0,9	0,9	0,96	0,96 starke Streuung
6370	0,9	0,9	0,96	0,98 starke Streuung
6970	0,94	0,98	1,0	0,8
8150	0,96	0,91	0,85	0,92
Mittel	0,98	0,90	0,95	0,96
U_i	10,2	10,2	10,2	10,15
U_k	9,1	9,2	9,7	9,75

Tabelle 4

Werte für $\frac{U_k}{U_i} \cdot p \sim 60 \text{ atm} = \text{const}$

α SKT	Wellenlänge λ	Leistungskonzentration kW/cm ³			
		54–66	66–85 Leitfähigkeit $\Omega^{-1} \text{ cm}^{-1}$	85–106	106–140
		1,8–2,15	2,15–2,6	2,6–3,1	3,1–3,7
1,2	3370	0,9	0,94	1,0	1,1
2,0	3440	0,92	0,92	1,04	1,25
3,4	3570	0,82	0,92	1,05	1,06
5,0	3750	0,98	1,05	1,07	1,1
	¹⁰⁾	0,78	0,92	1,1	1,2
0,8	3900	0,96	1,04	1,0	1,1
1,6	¹⁰⁾	0,96	0,98	0,98	1,1
2,3	3980	0,94	0,96	1,04	1,1
3,4	4500	0,98	1,05	1,05	1,1
	¹⁰⁾	0,87	0,92	1,02	1,1
1,1	4700	0,94	0,98	1,0	1,1
1,9	¹⁰⁾	0,96	0,98	1,1	1,3
3,1	5220	0,8	0,82	0,98	1,1
4,4	6380	0,92	0,88	0,87	1,1
	¹⁰⁾	0,92	0,94	0,98	1,1
0,9	6970	0,92	0,98	0,98	1,1
1,6	8150	0,88	0,96	1,0	1,3
2,6	Mittel . . .	0,91	0,96	1	1,1
3,7	U_i	10,15	10,1	10,0	9,9
	U_k	9,2	9,7		

Bei gleicher Leistung, gleicher Temperatur wird der Anteil der Stoßmolekülstrahlung, da die Intensität quadratisch mit dem Druck ansteigt, mit Erhöhung des Druckes größer werden.

Die nebenstehende Abb. 1 soll dieses Ergebnis veranschaulichen.

Beurteilung der Genauigkeit der Werte für U_k

Die gefundenen Werte für U_k sind von uns bestimmt worden auf Grund der Beziehung

$$S_k = C \cdot e^{-\frac{e U_k}{k T}} \quad (3)$$

Da die kontinuierliche Strahlung jedoch aus zwei Anteilen – Stoßmolekülstrahlung + Rekombinati-

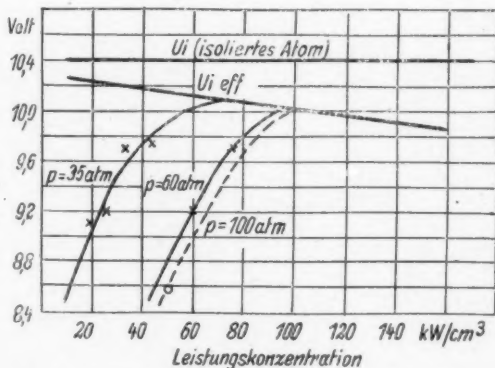


Abb. 1. „Anregungsspannung“ des Kontinuums in Abhängigkeit von der Leistung bei drei verschiedenen Drucken

¹⁰⁾ Messungen mit verkürzter Länge des Eintrittspaltes.

onsstrahlung — besteht, muß diese Beziehung genauer heißen:

$$S_k = C_M \cdot T^{-1,5} \cdot e^{-\frac{eU_M}{kT}} + C_R \cdot T \cdot e^{-\frac{eU_i}{kT}}. \quad (3a)$$

Nur wenn einer der beiden Summanden vernachlässigbar gering ist, wie es bei einem Druck von 35 atm und einer Leistungskonzentration von $\sim 50 \text{ kW/cm}^2$ und bei einem Druck von 60 atm und Leistungskonzentration $\sim 80 \text{ kW/cm}^2$ für die Stoßmolekülstrahlung der Fall ist, läßt sich das Verhältnis $\frac{U_k}{U_i}$ mit ausreichender Genauigkeit aus der Messung von Leitfähigkeit und Strahlungsintensität nach der oben beschriebenen Methode bestimmen.

In einer inzwischen erschienenen Arbeit kommt P. Schulz¹¹⁾ ebenfalls zu der Ansicht, daß die kontinuierliche Strahlung des Quecksilberhochdruckbogens sich aus Stoßmolekülstrahlung und Rekombinationsstrahlung zusammensetzt. Er weist hierbei darauf hin, daß sich das Intensitätsverhältnis beträchtlich zugunsten der Stoßmolekülstrahlung verschiebt, da die Rekombinationsstrahlung zum Teil stark reabsorbiert wird.

Diese und die vorhergehende Arbeit über die Größe der effektiven Ionisierungsspannung entstanden auf Anregung von Herrn Professor Rompe. Ihm und Herrn Professor Möglich, wie auch Herrn Professor P. Schulz danke ich vielmals für zahlreiche klärende Diskussionen.

¹¹⁾ G. Busz u. P. Schulz, Ann. Physik (6) 6, 232 (1949).

Berlin-Buch, Institut für Festkörperforschung der Deutschen Akademie der Wissenschaften zu Berlin.

(Bei der Redaktion eingegangen am 1. Februar 1950.)

(3a)

es bei
 N/cm^2
 N/cm^2

aus-
nsität

lls zu
ogens
zt. Er
unsten
n Teil

rungs-
a und
e ich

ie der

D
st

we
sch
gru
tra
die
eber
Int
In
sch
läß
wei
Ein

auf
netz
füh
mik
der
erfa

Ab
Ge

sch
ber

Met

zeich

bier

(6)

zeich

Δ

Die Minimumstrahlkennzeichnung als Mittel zur Genauigkeitssteigerung optischer Messungen und als methodisches Hilfsmittel zum Ersatz des Strahlbegriffes¹⁾

Von Hans Wolter

(Mit 7 Abbildungen)

Inhaltsübersicht

Die Kennzeichnung einer Ebene im Raum durch ein möglichst konzentriertes Lichtwellenbündel ist zwangsläufig ungenau; die Unschärfe gehorcht einem zur Heisenberg'schen Unschärferelation analogen Gesetz. Vorliegende Arbeit zeigt, wie diese Grenze grundsätzlich beliebig weit unterschritten werden kann, wenn auf die Energiekonzentration verzichtet wird. Der Gewinn ist in jedem Falle durch Lichtverlust zu erkaufen; dieser ist jedoch relativ am geringsten bei optimaler Kennzeichnung der Ebene als Null-ebene des Lichtes (genauer durch die Flanken zu ihren Seiten), die am besten durch Interferenz zweier gegenphasiger Lichtbündel hergestellt werden kann (Phasenverfahren). In vielen Fällen eignet sich auch die Interferenz mit Anordnungen, die ein Fresnel'sches Biprisma oder ähnliches verwenden (Verdoppelungsverfahren). Mit diesen Mitteln läßt sich die Unschärfe prinzipiell beliebig weit herabsetzen, praktisch in jedem Falle so weit, wie Störlicht (Streulicht, Inkohärenz u. ä.) oder Mängel der Optik es erlauben. Ein Gewinn an Meßgenauigkeit um den Faktor 25 gelingt mit einfachen Hilfsmitteln.

Die Anwendbarkeit dieser „Minimumstrahlkennzeichnung“ beschränkt sich nicht auf die Kennzeichnung einer Ebene oder unmittelbar verwandte Probleme (Strahlversetzung bei Totalreflexion, Lichtzeigergeräte u. ä.); die Minimumstrahlkennzeichnung führt auch zu einer Verbesserung der Schlierenverfahren, ist auf spektroskopische und mikroskopische Untersuchungen anwendbar und gibt ein Mittel, gewissen Fragen — wie der nach dem Weg des Lichtes — einen experimentell greifbaren und theoretisch streng erfaßbaren Sinn zu geben.

§ 1. Die Strahlunschärfe

Die Anwendung des Wortes „Strahl“ in der Physik schließt entweder eine Abstraktion von den Wellenerscheinungen ein, oder aber man denkt bei seinem Gebrauch an ein Wellenbündel, das in einem seiner Querschnitte die flächenhafte

¹⁾ Auszugsweise vorgetragen auf der Tagung der Deutschen Physikalischen Gesellschaft am 23. April 1949 in Hamburg. Über einige spezielle Anwendungen der Methode berichtete der Verfasser schon in folgenden Arbeiten:

a) Untersuchungen zur Strahlversetzung bei Totalreflexion des Lichtes mit der Methode der Minimumstrahlkennzeichnung, Z. Naturforschg. 5a, 143 (1950).

b) Verbesserung der abbildenden Schlierenverfahren durch Minimumstrahlkennzeichnung, Ann. Physik (6) 7, 182 (1950).

c) Zur Frage des Lichtweges bei der Totalreflexion, Z. Naturforschg. im Erscheinen.

d) Experimentelle und theoretische Untersuchungen zur Abbildung nichtabsorbierender Objekte, Ann. Physik (6) 7, 33 (1950).

e) Zur Abbildung zylindrischer Phasenobjekte elliptischen Querschnitts, Ann. Physik (6) 7, 147 (1950).

f) Zur Genauigkeitssteigerung optischer Messungen durch Minimumstrahlkennzeichnung, Z. Naturforschg. 5a, 130 (1950).

Ausdehnung $\delta f = \delta x \cdot \delta y$ und die Winkelstreuung $\delta \omega = \delta x_x \cdot \delta x_y$ — in naheliegenden Bezeichnungen — hat. Ist λ die Wellenlänge, so zeigt die Unschärfebedingung

$$\delta x \cdot \delta x_x \geq \lambda; \delta y \cdot \delta x_y \geq \lambda, \quad (1)$$

die der Heisenbergschen analog ist, daß Winkelstreuung und Strahlbreite nicht zugleich beliebig klein gemacht werden können. Ungleichung (1) gilt in dieser Form nur für kleine Winkel, bei größeren hat der Sinus an die Stelle des Winkels selbst zu treten, und der Winkel ist dann von einer im Einzelfall noch zu definierenden Symmetrieachse aus zu zählen. Führt man die praktische „Richtungsvariable“

$$\gamma = \frac{\sin \alpha}{\lambda} \quad (2)$$

ein, so nimmt die Unschärfebedingung die Form an $\delta x \cdot \delta \gamma \geq 1$; die reine Zahl $\delta x \cdot \delta \gamma$ bezeichnen wir als „Unschärfezahl“ oder kurz als „Unschärfe“, das Reziproke als „Schärfe“²⁾.

Die für jede Strahlung gültige Unschärfebedingung hat speziell in der Optik die bekannten Auswirkungen auf das Auflösungsvermögen abbildender optischer Instrumente. Bei der Messung von Strahlversetzungen, wozu in einer früheren Arbeit¹⁾ ein Beispiel behandelt wurde, und bei der Messung von Strahlswenkungen — z. B. bei Lichtzeigegeräten wie Spiegelgalvanometern — begrenzt sie die Meßgenauigkeit.

Eine Unterschreitung jener Grenze ist daher oft erwünscht; sie erscheint aber grundsätzlich unmöglich, solange man als Zeiger ein solches Maximum des Lichtes verwendet, das den Hauptteil der Energie der Welle führt. Denn für diesen Teil der Welle folgt die Unschärfebedingung völlig zwangsläufig aus der Wellengleichung. Mit der Erkenntnis über die Voraussetzungen der Unschärfebedingung eröffnen sich aber sofort drei Wege zu ihrer Umgehung bei dem Lichtzeiger und verwandten Problemen. Erstens kann man versuchen, als Kennzeichnung einer Ebene mittels des Lichtes ein solches Maximum zu verwenden, das nur einen Teil der Energie führt, zweitens kann man statt dessen eine möglichst steile Flanke oder drittens ein Minimum, besser eine „Nullebene“ der Intensität heranziehen.

Alle drei Wege wurden im Rahmen dieser Arbeit experimentell untersucht; die beiden ersten wurden dabei aber von dem dritten so weit übertroffen, daß sie nur zum Vergleich und zum Nachweis ihrer Unterlegenheit diskutiert werden sollen.

§ 2. Auswirkung einer Lichtverteilung bei visueller oder photographischer Auswertung

Unsere — das Ergebnis der folgenden Untersuchungen vorwegnehmende — Behauptung, daß Minima die bei weitem schärfste Strahlkennzeichnung ermöglichen, forderte mehrfach den Einwand heraus, daß die Intensität als Funktion der Richtvariablen Maxima und Minima in gleicher Breite aufzuweisen pflegt. Dem ist zu entgegnen, daß es in der Praxis auf Intensitätsverhältnisse und nicht auf Intensitätsdifferenzen ankommt.

²⁾ Der Einfachheit halber interessieren wir uns hier nur für eine der beiden Dimensionen senkrecht zur Fortpflanzungsrichtung.

³⁾ Daß diese Definition über unsere Fragestellung hinaus zweckmäßig ist, erhellt aus einem bekannten Satz der geometrischen Optik, der mit diesem Begriff die einfache Form annimmt: Ideale abbildende Mittel (Linsen, Spiegel, Prismen und beliebige Kombinationen aus solchen) lassen die Unschärfezahl unverändert (Sinusbedingung).

Für die unmittelbare visuelle Beobachtung ist das der Inhalt des Weber-Fechner'schen Gesetzes, das man in seiner integrierten Form wie folgt auszusprechen pflegt: Die Stärke der Lichtempfindung ist dem Logarithmus der Lichtstärke proportional — und wir können ergänzen: dann auch dem Logarithmus der Feldstärke proportional. (Über eine durch den Schwellenwert des Auges bedingte Grenze wird unten im Zusammenhang mit den Experimenten noch die Rede sein.) Diese für die Beurteilung einer Lichtverteilung wichtige Feststellung läßt sich auch auf photographische Auswertung ausdehnen; dabei soll zugleich der Begriff der Strahlschärfe eine experimentell bequem erfassbare und präzise Form erhalten.

Unter der „Graubreite“ einer Hell-Dunkel-Grenze in einer Aufnahme verstehen wir den Abstand $\xi_h - \xi_d$ einer im Positiv gerade weiß erscheinenden Stelle von der benachbarten dunklen Grenze, die bei der Aufnahme gerade den Schwellenwert der Belichtung erfahren habe. Bezeichnen wir die Belichtungszeit mit t , den Schwarzschildexponenten mit p und die Feldstärke mit E , die Beleuchtung selbst mit $C \cdot E^2$, so sei $C \cdot E^2 \cdot t^p \cdot e^{-w}$ der Schwellenwert der Belichtung und $C \cdot E^2 \cdot t^p \cdot e^w$ die Vollbelichtung, d. h. die Belichtung, die zu der auf dem Positiv weißen Grenze des Graubereichs führt. $C \cdot E^2 \cdot t^p = G$ heiße die Graubelichtung und w die Weichheit des photographischen Verfahrens. Für eine scharfe Strahlkennzeichnung wird man praktisch möglichst kleines w anstreben, also große Härte des Aufnahmемaterials. Bei den z. B. den Abb. 2 zugrundeliegenden Verhältnissen war $w = 0,3$ (IG-Stat-Film vergrößert auf Agfa-Lupex-Hart).

Entwickelt man die Belichtung als Funktion der Koordinate auf der Aufnahme ξ in die Taylorsche Reihe und bricht nach dem ersten Gliede ab, so erhält man für die Graubreite

$$\xi_h - \xi_d = \frac{C \cdot E^2 \cdot t^p \cdot e^w - C \cdot E^2 \cdot t^p \cdot e^{-w}}{\frac{d}{d\xi} C \cdot E^2(\xi) \cdot t^p} = \frac{\sin w}{\frac{d}{d\xi} \ln |E(\xi)|} \quad (3)$$

sofern der im Nenner stehende Differentialquotient nicht verschwindet oder zu klein wird, um noch das Abbrechen der Reihe zu rechtfertigen. Diesen Fall wollen wir zunächst ausschließen und später gesondert betrachten. Wir befassen uns also zunächst nicht mit Stellen im Richtdiagramm, in denen ein Extremwert vorliegt, oder mit einer sehr weichen photographischen Abbildung, die ja ohnehin kaum interessiert. Der Nenner ist an einer Stelle innerhalb der Graubreite zu nehmen; wir ziehen dazu den Grauwert ξ_0 heran, in dem die Graubelichtung vorliegt.

Die Graubreite wird also stets durch den Feldstärkenlogarithmus diktiert, unabhängig davon, ob visuell oder photographisch ausgewertet wird. Sie sei nach (3) für eine Flanke und anschließend für Extremwerte berechnet.

In der Umgebung einer Flankenstelle ξ_0 sei der Verlauf der Feldstärke durch

$$|E(\xi)| = |E(\xi_0)| + s \cdot (\xi - \xi_0) \quad (4)$$

gegeben. Nach (3) folgt für die Graubreite

$$\xi_h - \xi_d = \frac{\sin w \cdot E(\xi_0)}{s}; \quad (5)$$

für eine möglichst scharfe Strahlkennzeichnung ist also selbstverständlich große Flankensteilheit zu fordern.

An einem Extremwert, an dem die Feldstärke nicht den Wert 0 annimmt, sei ihr Betrag durch die Reihenentwicklung

$$|E(\xi)| = |E(\xi_0)| \cdot (1 + \frac{1}{2} \cdot K \cdot (\xi - \xi_0)^2 + \dots) \quad (6)$$

gegeben. K werde als relative Krümmung bezeichnet. An einem Minimum ist K positiv. Durch Wahl der Belichtungszeit sei die dunkle Grenze des Graubereichs in den Scheitel des Minimums gerückt $\xi_0 = \xi_d$; dann ist $E(\xi_h) = E(\xi_d) \cdot e^w$ und $E(\xi_h) = E(\xi_d) \cdot (1 + \frac{1}{2} \cdot K \cdot (\xi_h - \xi_d)^2)$. Daraus folgt die Graubreite

$$\xi_h - \xi_d = \pm 2 \cdot \sqrt{\frac{\frac{w}{2} \cdot \sin \frac{w}{2}}{K}} \quad (7)$$

eines Minimums. Wenn wir bei einem Maximum die helle Grenze des Graubereichs in den Scheitel legen, so folgt entsprechend die Graubreite

$$\xi_h - \xi_d = \pm 2 \cdot \sqrt{\frac{\frac{w}{2} \cdot \sin \frac{w}{2}}{|K|}} \quad (8)$$

eines Maximums in nur unwesentlicher Abweichung von (7). Bei einer Strahlkennzeichnung durch Extremwerte ist also große relative Krümmung im Scheitel anzustreben.

Diese einfachen und sehr allgemeinen Ergebnisse sollen nun auf Richtcharakteristiken übertragen werden. Zu dem Zweck denken wir uns das Licht einer zu untersuchenden strahlenden Anordnung durch eine ideale Sammellinse in deren Brennebene, die als Beobachtungsebene dient, geworfen. Jeder Punkt dieser Ebene repräsentiert dann einen Einfallswinkel, d. h. eine Richtvariable $\gamma = \frac{\sin \alpha}{\lambda}$. Als Koordinate ξ in der Beobachtungsebene wählen wir γ selbst, und nennen die komplexe Feldstärke $E(\gamma)$ als Funktion der Richtvariablen die Richtcharakteristik der strahlenden Anordnung. Richtdiagramm heiße die als ebene Kurve darstellbare Funktion $|E(\gamma)|$.

Zur Anwendung der Ergebnisse dieses Paragraphen auf Richtcharakteristiken haben wir lediglich überall die Koordinate ξ und die Richtvariable γ miteinander zu identifizieren.

Die so definierte Graubreite $\gamma_h - \gamma_d$ führt im Spezialfall $w = \ln 2$ bei einem Maximum oder Minimum auch auf die Halbwertsbreite im üblichen Sinne $\delta\gamma = 2 \cdot (\gamma_h - \gamma_d)$; dieser Wert ist zur Berechnung der Unschärfenzahl $\delta x \cdot \delta\gamma$ zu verwenden.

§ 3. Der durch eine Querschnittsbegrenzung (Aperturblende) auf die Richtcharakteristik ausgeübte Zwang

Wie die Unschärfebedingung zeigt, könnte die Breite eines Wellenbündels auf einem Schirm δx die Mindestbreite von der Größenordnung einer Wellenlänge nur dann annehmen, wenn die Apertur des auf den Schirm fallenden Bündels nicht wesentlich begrenzt wäre. In den hier zur Diskussion stehenden Problemen sollen jedoch Querschnitt und Apertur zugleich klein gehalten werden, bzw. die Vorrichtung, in der ein „Strahl“ zur Anwendung kommt, hat zwangsläufig eine endliche Durchlaßblende. Als solche ist z. B. bei den Drehspiegelgeräten der Spiegel selbst wirksam, bei dem Spektroskop eine Prismenfläche und im Mikroskop die Aperturblende. Da diese Durchlaßblende für das auf die Beobachtungsebene auffallende Licht als Aperturblende wirksam ist, soll sie im folgenden schlechthin als Aperturblende bezeichnet werden (im Unterschied zu der in der Beobachtungsebene selbst liegenden Gesichtsfeldblende). Diese Bezeichnung möge nicht die einfache Tatsache verschleiern, daß es sich um eine Querschnittsbegrenzung der Welle handelt an einem Orte, der von der Beobachtungsebene als hinreichend weit entfernt gelten kann.

Flankensteilheit und Krümmung des Richtdiagramms und darüber hinaus auch die höheren Ableitungen in der Richtcharakteristik werden nun in ganz charakteristischer Weise durch die Größe dieser Aperturblende beschränkt. Das sei Gegenstand der anschließenden Untersuchung; wir werden, um die Fragestellung zu präzisieren, dabei die Aperturblende als gegeben ansehen und nach dem Feldstärkenverlauf innerhalb der Aperturblende fragen, den wir herstellen müssen, um optimale Strahlkennzeichnung zu erreichen.

Die Feldstärkenverteilung innerhalb der Aperturblende sei durch

$$A \cdot B(x) \cdot e^{i\omega t}$$

beschrieben. A sei eine positive Konstante, die wir z. B. zur Erfassung der Lichtquellenlichtstärke verfügbar halten; das als „Belegungsfunktion“ bezeichnete

$B(x)$ sei Null für $|x| > \frac{b}{2}$ — das beschreibt die Wirkung der Aperturblende — und möge sonst beliebig komplex zugelassen werden. Praktisch ist es, $|B(x)| \leq 1$ bereits von vornherein festzusetzen; das geschieht wegen des verfügbaren A ohne Beschränkung der Allgemeinheit und läßt die zweckmäßige Deutung zu, daß A die über die ganze Blende gleichmäßige Amplitude einer einfallenden Welle ist und $B(x)$ den Einfluß einer dort angebrachten — Amplitude und Phase verändernden Platte darstellt. $|B(x)| \leq 1$ besagt in diesem Bilde dann, daß jene Platte die Amplitude natürlich nicht verstärken, sondern höchstens schwächen kann. Die Betrachtungen sind aber selbstverständlich völlig unabhängig davon, in welcher Weise eine Belegung zustande gekommen ist und gelten daher viel allgemeiner als nur in solchen Fällen, auf die ein solches Bild anwendbar ist. Ebenso ist die Gültigkeit der Ergebnisse nicht an die reale Existenz einer Aperturblende gebunden; es genügt vielmehr, daß ein Streifen endlicher Breite b existiert, außerhalb dessen die Feldstärke verschwindet.

Alle Quellen mögen in großer Entfernung auf der einen, die Beobachtungsebene ebenfalls in großer Entfernung auf der anderen Seite der Blende liegen. Da außerdem nur scharfe Strahlkennzeichnung interessiert, betrachten wir die Richtcharakteristik nur in der Umgebung der Blendennormalen. Würde statt der Normalen eine zur Blende schräg liegende Gerade ausgezeichnet, so wäre durch alle Formeln ein Phasenfaktor zu schleppen; die wesentlichen Ergebnisse blieben unverändert. Unter diesen Voraussetzungen folgt aus der entsprechend vereinfachten Kirchhoffschen Formel unter Fortlassen eines Proportionalitätsfaktors und des Zeitfaktors $e^{i\omega t}$ die Richtcharakteristik

$$E(\gamma) = A \cdot \int_{-\infty}^{\infty} B(x) \cdot e^{-2\pi i x \gamma} dx = A \cdot \sum_{\nu=0}^{\infty} \frac{(-2\pi i)^{\nu}}{\nu!} \cdot M_{\nu} \cdot \gamma^{\nu} \quad (9)$$

mit

$$M_{\nu} = \int_{-\infty}^{\infty} x^{\nu} \cdot B(x) dx; \quad \gamma = \frac{\sin \alpha}{\lambda}.$$

Die Momentintegrale 1. und 2. Ordnung über die Belegungsfunktion lassen daher einen Schluß darüber zu, in welcher Weise die endliche Größe der Aperturblende einen Zwang auf die Flankensteilheit und die Krümmung ausübt; und zwar folgt aus $|B(x)| \leq 1$ und den Symmetrieeigenschaften der Funktion x^{ν} die Ab-

schätzung

$$M_v = \int_0^{\frac{b}{2}} x^v \cdot \{B(x) + (-1)^v \cdot B(-x)\} dx \quad (10)$$

$$|M_v| \leq \left| 2 \int_0^{\frac{b}{2}} x^v dx \right| = \frac{b^{v+1}}{(v+1) \cdot 2^v}. \quad (11)$$

An die Stelle des \leq tritt das Gleichheitszeichen dann und nur dann, wenn

$$|B(x)| = 1; \quad B(x) = \text{konstant für } x > 0; \quad B(-x) = (-1)^v \cdot B(x). \quad (12)$$

Die physikalische Auswirkung dieser zunächst mathematischen Erkenntnis wird für Maxima, Flanken und Minima getrennt untersucht.

§ 4. Die Maximumverschärfung

Vernachlässigt man in der Richtcharakteristik (9) Glieder 3. und höherer Ordnung in γ und betrachtet den Fall eines Extremums, so wird das Richtdiagramm durch

$$|E(\gamma)| = |E(0)| \cdot (1 + \frac{1}{2} \cdot K \cdot \gamma^2) \quad (13)$$

dargestellt und kann aus der Richtcharakteristik (9) durch Bilden des Betrages zu

$$|E(\gamma)| = A \cdot |M_0| \cdot \left\{ 1 + \gamma^2 \cdot 2\pi^2 \cdot \left(Re^2 \frac{M_1}{M_0} - Re \frac{M_2}{M_0} \right) \right\} \quad (14)$$

gefunden werden; man erhält also für die relative Krümmung

$$K = 4\pi^2 \cdot \left\{ Re^2 \frac{M_1}{M_0} - Re \frac{M_2}{M_0} \right\}. \quad (15)$$

In diesem Paragraphen seien weiterhin nur Maxima untersucht; da der erste Summand in (15) stets positiv ist, wird dann unter Benutzung von (11)

$$-K = |K| \leq 4\pi^2 \cdot \left| Re \frac{M_2}{M_0} \right| \leq 4\pi^2 \cdot \left| \frac{M_2}{M_0} \right| \leq \frac{\pi^2}{3} \cdot \frac{b^3}{|M_0|}. \quad (16)$$

M_0 hat bei gleichmäßiger und möglichst stark belegter Aperturblende ($B(x) \equiv 1$) selbst seinen Maximalwert b ; die Bedingung dafür (12) ist erfüllt. Da in diesem Falle in (16) überall das Gleichheitszeichen gilt, ist dann $-K = \frac{\pi^2}{3} \cdot b^2$. Gegenüber diesem Spezial-

fall, der in der üblichen Form des Lichtzeigers realisiert ist, kann bei gleicher Aperturblende die relative Krümmung nach (16) nur dadurch erhöht werden, daß eine Belegung der Aperturblende erfolgt, die ein kleineres M_0 zur Folge hat. Nun ist nach (9) aber $|A \cdot M_0|$ die Feldstärke im Maximum selbst; schärfere Strahlkennzeichnung ist also automatisch mit verringerter Feldstärke im Maximum gekoppelt.

Wird eine Abnahme der Feldstärke im Scheitel gegenüber dem maximal möglichen Wert $A \cdot b$ auf $m \cdot A \cdot b$ (mit $0 < m \leq 1$) zugelassen, so ist $M_0 = m \cdot b$, und die Schranke für die relative Krümmung kann auf

$$|K| \leq \frac{\pi^2}{3} \cdot b^2 \cdot \frac{1}{m} \quad (17)$$

heraufgesetzt werden. Die Gratbreite des Maximums wird daher nach Gl. (8)

$$(\gamma_h - \gamma_d)_m \geq \pm \frac{2}{\pi} \cdot \sqrt{3 \cdot e^{-\frac{w}{2}} \cdot \sin \frac{w}{2} \cdot \sqrt{m}}. \quad (18)$$

Wir wollen diese mit der Graubreite $(\gamma_h - \gamma_d)_1$ einer üblichen Maximumstrahlkennzeichnung (gleichmäßige Belegung $B \equiv 1$, $m = 1$) vergleichen und bezeichnen als Gewinn der

verschärften Maximumstrahlkennzeichnung das Verhältnis

$$g = \frac{(\gamma_h - \gamma_d)_1}{(\gamma_h - \gamma_d)_m}; \quad (19)$$

Da für $m = 1$ in der Abschätzung (18) das Gleichheitszeichen gilt (siehe auch die Betrachtung im Anschluß an Gl. (16)), so erhält man für den Gewinn

$$g \leq \sqrt{\frac{1}{m}}. \quad (20)$$

Der Gewinn muß also — wenn er überhaupt möglich ist — mindestens unter Intensitätsverlust, d. h. mit längerer Belichtungszeit T erkauft werden gegenüber der Belichtungszeit t des üblichen Verfahrens ($m = 1$); denn es gilt $m^2 \cdot T^p = t^p$ und nach (20) also

$$g \leq \tau^{\frac{1}{p}} \quad \text{mit} \quad \tau = \left(\frac{T}{t}\right)^p. \quad (21)$$

Bei einem Schwarzschildexponenten p nahe 1 erfordert ein Gewinn von einer Größenordnung eine Erhöhung der Belichtungszeit um mindestens 4 Größenordnungen. Wir werden sehen, daß deshalb dieses Maximumverfahren gegenüber den in dem nächsten Paragraphen zu behandelnden Flanken- und Minimumverfahren nicht konkurrenzfähig ist.

Beispiele zur Realisierung der Maximumverschärfung und Aufnahmen hierzu seien deshalb in diesen Auszug nicht aufgenommen. Die Richtdiagramme zeigen relativ scharfe und dabei naturgemäß schwache Maxima, die durch Nullstellen von wesentlich helleren Maxima getrennt werden.

§ 5. Flanken kennzeichnung und Minimumkennzeichnung

Man sagt, daß die Richtung $\gamma = 0$ an einer „Flanke“ der Richtcharakteristik liegt, wenn in dieser

$$E(\gamma) = A \cdot (M_0 - 2\pi i M_1 \cdot \gamma + \dots) \quad (22)$$

die Glieder 2. und höherer Ordnung neben dem 1. Ordnung in einer durch die photographische Weichheit w bestimmten Umgebung von $\gamma = 0$ vernachlässigbar sind. Durch Bilden des Betrages findet man dann leicht aus der Richtcharakteristik das Richtdiagramm

$$|E(\gamma)| = A \cdot \left(|M_0| + 2\pi \gamma \cdot \operatorname{Im} \frac{M_1 \cdot |M_0|}{M_0} \right) \quad (23)$$

und die Flankensteilheit

$$s = A \cdot 2 \cdot \pi \operatorname{Im} \frac{M_1 \cdot |M_0|}{M_0} \quad \text{für } M_0 \neq 0 \quad \text{bzw.} \quad |s| = A \cdot 2 |M_1| \cdot \pi \quad \text{für } M_0 = 0. \quad (24)$$

Da nach (5) die Graubreite bei gleicher Belichtungszeit und gleichem Photomaterial um so kleiner ist, je größer die Flankensteilheit dem Betrage nach ist, wünschen wir $|s|$ möglichst zu steigern; doch ist offenbar nach (24)

$$|s| \leq A \cdot 2\pi \cdot |M_1| \quad \text{für } M_0 \neq 0 \quad \text{bzw.} \quad |s| = A \cdot 2\pi \cdot |M_1| \quad \text{für } M_0 = 0, \quad (25)$$

und der durch die Aperturblende ausgeübte Zwang (11) setzt der Flankensteilheit deshalb die Schranke

$$|s| \leq A \cdot \frac{\pi}{2} \cdot b^2. \quad (26)$$

Optimal sind Belegungen $B(x)$, für die statt der \leq in (25) und (11) für $\nu = 1$ das = tritt. Bezüglich des zweiten Teiles der Abschätzung ist das nach (12) der Fall, wenn

$$B(x) = \begin{cases} 1 & \text{für } x > 0 \\ -1 & \text{für } x < 0. \end{cases} \quad (27)$$

Diese Belegung ist aber auch im wesentlichen die einzige, die diese Bedingung erfüllt (nämlich abgesehen von einem physikalisch uninteressanten konstanten Phasenfaktor und abgesehen von Abänderungen an einer Argumentmenge von Maß 0). Der erste Teil der Abschätzung, die Ungleichung (25), verwandelt sich dann automatisch ebenfalls in eine Gleichheit, da hier

$$M_0 = \int B(x) dx = 0 \quad (28)$$

ist. Nach Gl. (24) ist

$$|s| = A \cdot 2\pi \cdot |M_1| = A \cdot 2\pi \cdot \int x \cdot B(x) dx = A \cdot \frac{\pi}{2} \cdot b^2 \quad (29)$$

und nimmt also den Maximalwert der Flankensteilheit überhaupt an.

Das Bemerkenswerteste an dieser Lösung ist die Tatsache, daß — wie mit (28) bereits ausgesprochen — die Forderung größter Flankensteilheit an $\gamma = 0$ zwangsläufig zu $M_0 = 0$, d. h. $E(0) = 0$ also zu einem Intensitätsminimum an $\gamma = 0$ führt. Die Reihenentwicklung der Richtcharakteristik beginnt dann mit

$$E(\gamma) = s \cdot \gamma + \dots,$$

und das Richtdiagramm

$$|E(\gamma)| = |s| \cdot |\gamma| + \dots$$

hat an $\gamma = 0$ ein „lineares Minimum“, wie es im Unterschied zu einem „quadratischen Minimum“ der Richtcharakteristik

$$E(\gamma) = E(0) + 2\pi^2 \cdot |M_2 \cdot A| \gamma^2 + \dots \quad (30)$$

im Anschluß an den üblichen Sprachgebrauch bezeichnet sei.

In die Nullstelle selbst kann man den Schwellenwert des Aufnahmемaterials natürlich durch keine noch so lange Belichtungszeit rücken (wir sehen hier von Streulicht und anderen Störungen zunächst ab, werden seine modifizierende Wirkung jedoch noch unten betrachten); an der Flanke eines linearen Minimums $|E(\gamma)| = |s| \cdot |\gamma|$ in solcher Nähe der Nullstelle, daß die Flankensteilheit noch unverändert ist, berechnet sich die Graubreite zu

$$(\gamma_h - \gamma_d)_n = \pm 2 \cdot \sin \frac{w}{2} \cdot \frac{E(\gamma_s)}{|s|} = \pm 2 \cdot \sin \frac{w}{2} \cdot \frac{1}{|s|} \sqrt{\frac{G}{C}} \cdot T^{\frac{p}{2}}. \quad (31)$$

Für $|s|$ ist bei der optimalen Belegung (27) der Wert $A \cdot \frac{\pi}{2} \cdot b^2$ nach (29) zu verwenden. Setzen wir die damit erhaltene Graubreite ins Verhältnis zu der Graubreite einer gleichmäßig belegten Aperturblende [Gl. (18) für $m = 1$], so erhalten wir den Gewinn

$$g = \frac{1}{2} \sqrt[3]{3 \cdot \frac{\tau^{\frac{1}{2}}}{\sqrt{e^{\frac{w}{2}} \cdot \sin \frac{w}{2}}}} \quad \text{mit} \quad \tau = \left(\frac{T}{t}\right)^p. \quad (32)$$

Je länger belichtet wird, je näher die Graugrenze also dem Minimum rückt, desto schmaler wird auch der Graubereich. Der Gewinn muß auch hier durch längere Belichtungszeit erkauft werden; aber der größere Exponent bei τ gegenüber dem Exponenten $\frac{1}{2}$, der bei der Maximumstrahlkennzeichnung auftrat, ist wesentlich günstiger. Zu einem Gewinn, der dort eine Verlängerung der Belichtungszeit um den Faktor 100 erfordern würde, ist hier der Faktor 10 ausreichend. Das gilt zunächst bei einer Weichheit w des photographischen Verfahrens, die nicht wesentlich unter 1 liegt; doch hat die Minimummethode ersichtlich von einem weiteren

g er-
nten
von
sich

(28)

(29)

(28)
ange-
= 0

adra-

(30)

erials
von
Wir-
nums
h un-

(31)

ver-
Grau-
alten

(32)

desto
ngere
dem
ntlich
t um
lt zu-
esent-
teren

F
u
L

m
is
cl

ur

di
de

M
op

nä
Fr

Ab

1.
zu
We
ode
erst
knü
met
nur

ein
übe

Herabsetzen der Weichheit w einen größeren Vorteil als die Maximummethode, und der Gewinn läßt sich dadurch entsprechend der Gl. (32) wesentlich steigern. Diese Dinge werden im einzelnen in den nächsten Paragraphen betrachtet werden.

Ein quadratisches Minimum

$$E(\gamma) = \frac{1}{2} \cdot \kappa \cdot \gamma^2 \quad (33)$$

mit der absoluten Scheitelkrümmung κ bietet, wie anschaulich ohne weiteres zu erwarten ist, nicht die Vorteile eines linearen Minimums. Der von der Aperturblende auf die Richtcharakteristik ausgeübte Zwang führt ähnlich wie oben bei den Maxima auf

$$\kappa \leq \frac{\pi^2}{3} \cdot b^2 \cdot A$$

und gegenüber der gleichmäßig belegten Aperturblende auf den Gewinn

$$g \leq \sqrt{2} \cdot \frac{\tau^{\frac{1}{2}}}{\sqrt{\sin \frac{w}{2}}} ; \quad (34)$$

die Abhängigkeit von der photographischen Weichheit w ähnelt der bei den linearen Minima, der Gang mit dem Belichtungsverhältnis τ dem bei den Maxima.

Im folgenden werden ausschließlich die als überlegen erwiesenen linearen Minima betrachtet, und zwar zunächst die nach den bisherigen Untersuchungen optimale Realisierung mit der gegenphasigen Belegung der Aperturblende (27).

§ 6. Das Phasenverfahren bei Abbildung auf eine feste Ebene

Bisher war kurz von Strahlkennzeichnung die Rede; man muß jedoch bei näherer Betrachtung zwei Aufgaben unterscheiden, deren Lösungen ähnlich wie Fraunhofersche und Fresnelsche Beugungsverfahren zueinander stehen.

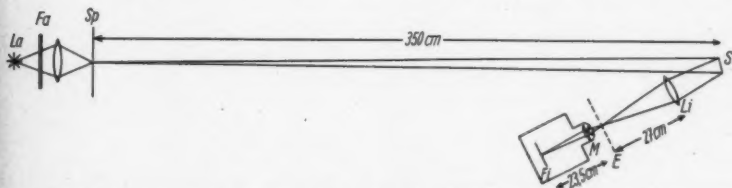


Abb. 1. Lichtzeigeranordnung für Drehspiegelgeräte als Beispiel für eine Vorrichtung mit Lichtmarke in fester Ebene

1. In vielen Fällen genügt es, auf einer festen Skala eine bewegliche Lichtmarke zu beobachten. 2. In anderen Fällen, z. B. solchen, in denen das Licht auf seinem Wege verfolgt werden soll, muß die Strahlkennzeichnung auf dem ganzen Wege oder doch einem größeren Teile desselben optimale Schärfe haben. Zuerst sei der erste Fall untersucht; um etwas Bestimmtes und Einfaches vor Augen zu haben, knüpfen wir mit Abb. 1 an ein Lichtzeigergerät nach Art eines Spiegelgalvanometers an, ohne uns auf solche speziellen Anwendungen festzulegen oder sie auch nur als besonders wesentliche zu bezeichnen.

Ein von der Glühlampe La über das (nicht unbedingt hochwertige und nicht einmal immer notwendige) Farbfilter Fa rückwärts beleuchteter Spalt Sp wird über den Drehspiegel S des Meßsystems z. B. eines Spiegelgalvanometers durch

die Objektivlinse Li auf die Ebene E abgebildet; das Bild kann wahlweise mit einer Lupe auf einer Skala beobachtet oder durch ein Mikroskopobjektiv M , dessen numerische Apertur größer als die Apertur des Strahlenbündels ist, auf einen Film Fi abgebildet werden. Eine praktische Ausführung, mit der die Aufnahmen zu den Abb. 2 auf IG-Stat-Film hergestellt wurden, hatte die in Abb. 1 eingetragenen Abmessungen. Die Mikrokamera vergrößerte 11,4fach, die Aufnahmen wurden 9,2fach nachvergrößert, so daß die Gesamtvergrößerung 105fach gegenüber dem Bild in der Ebene E betrug. Bei den hier ausgewählten Aufnahmen war statt einer Glühlampe mit Filter eine Quecksilberlampe mit Ausfilterung der grünen Linie ($\lambda = 0,546 \mu$) benutzt worden. Die Abb. 2a und 2b wurden so bei verschiedenen Belichtungszeiten (0,5 sec bzw. 5 sec) hergestellt und zeigen ersichtlich die bekannte Lichtmarke der Lichtzeigergeräte, nämlich die Erscheinung der Fraunhoferschen Beugung am Spalt; als solcher ist der kleine Drehspiegel (hier rechteckig) von 4 mm wirksamer Breite anzusehen. Setzt man die „Breite“ des Mittelmaximums im Sinne der Unschärfebedingung gleich 4 mm in der Abbildung, so berechnet sich nach den oben angegebenen Vergrößerungen in der Ebene E eine Querschnittsbreite des Lichtbündels $\delta x = 0,0382$ mm; die Winkelstreuung in dieser Ebene berechnet sich aus der Spiegelbreite (4 mm) und dem Abstand der Objektivlinse von der Ebene E (27 cm) zu $\delta x_x = 0,0149$. Mithin ist $\delta x_x \cdot \delta x = 0,57 \mu = 1,04 \cdot \lambda$. Die Unschärfezahl 1,04 ist also etwa die kleinste, die nach der Unschärfebedingung (1) für ein Maximum, das wie das hier abgebildete den Hauptteil der Energie führt, überhaupt möglich ist.

Abb. 2c, 2d und 2e sind Lichtmarken, bei denen diese Unschärfezahl mit dem in den vorhergehenden Paragraphen begründeten Minimumverfahren herabgesetzt wurde; und zwar wurde die Aperturblendenbelegung (27), d. h. Gegenphasigkeit der Blendenhälften, dadurch erreicht, daß für die Aufnahmen 2c und 2d nahe an dem als Aperturblende wirkenden Drehspiegelchen S eine Phasenplatte eingefügt wurde. Diese verschob das Licht der einen Spiegelhälfte in Phase um 180° gegenüber dem Licht der anderen Hälfte.

Eine solche Phasenplatte ist z. B. eine Planparallelplatte, deren eine Hälfte mit einem dünnen Zaponlacküberzug geeigneter Dicke bedeckt ist. Eine im Hochvakuum aufgedampfte Quarzschicht ist schwieriger aber gleichmäßiger herzustellen. Einfacher, vielseitiger und besser ist ein „Phasenschieber“ aus zwei Planparallelplatten, die mit einer Kante aneinanderstoßend gemeinsam so auf einen Rahmen gekittet werden, daß ihre Ebenen einen Winkel von fast 180° miteinander bilden. Das Plattenpaar wird dann so montiert, daß es gemeinsam um eine in der Trennkante beider Platten liegende vertikale Achse drehbar ist. Durch diese Drehung kann die Phasendifferenz φ der beiden Hälften eines hindurchtretenden Wellenbündels verändert werden; sie ist für kleine Winkel eine annähernd lineare Funktion des Drehwinkels α , und zwar bei entsprechender Festsetzung der Nullrichtung

$$\varphi = \alpha \cdot \frac{n-1}{n} \cdot \frac{D}{\lambda} \cdot \varepsilon, \quad (35)$$

wenn $180^\circ - \varepsilon$ der Winkel zwischen beiden Platten, n der Brechungsindex und D die Dicke der Platten ist. Die Eichung geschieht zweckmäßig empirisch. Für die Aufnahmen 2c und 2d war die Phasenverschiebung auf 180° eingestellt.

Alle diese Lösungen bedingen aber eine schwierige Justierung, deshalb wurde ein „Phasenspiegel“ entwickelt, der als Drehspiegel den Lichtzeigergeräten selbst einzusetzen ist und dann eine Benutzung des Instruments in der normalen Anordnung ermöglicht; zugleich werden Streulicht verursachende Flächen eingespart. Der Phasenspiegel kann durch Überziehen eines üblichen Galvanometerspiegels zur Hälfte mit einer Lack- oder Quarzschicht passender Dicke hergestellt werden.

Besonders bewährt hat sich ein Verfahren, bei dem eine ebene Glasfläche im Hochvakuum zunächst zur Hälfte mit einer $\lambda/4$ dicken Silberschicht und anschließend als Ganzes mit einer dickeren Silberschicht bedampft wird.

Die mit dem Phasenschieber hergestellten Aufnahmen 2c und 2d unterscheiden sich voneinander nur durch die Belichtungszeit (3 sec bzw. 50 sec). Sie war bei 2d so gewählt, daß innerhalb des Minimums schon das Streulicht den Schwellenwert des Films überschritt; die wirksame Breite des Minimums beträgt 0,3 mm und verhält sich zu der Halbwertsbreite des ungekennzeichneten Maximums wie 1:15.

Abb. 2e ist mit derselben Apparatur, jedoch mit einem Phasenspiegel statt des Phasenschiebers und mit 120facher statt 105facher Gesamtvergrößerung hergestellt worden. Die Vergrößerung wurde erhöht, soweit das ohne Wechsel der Optik möglich war, da sich das kennzeichnende Minimum in diesem Falle als überraschend scharf herausgestellt hatte. Die Minimumbreite beträgt in dieser Aufnahme 0,2 mm und verhält sich bei Berücksichtigung der etwas verschiedenen Vergrößerung zur Breite des Maximums eines ungekennzeichneten Lichtzeigers (Abb. 2a) wie 1:26. Die größere Schärfe gegenüber dem mit Phasenschieber gewonnen Minimum beruht größtenteils auf den einwandfreien ebenen Flächen des Phasenspiegels.

Setzt man die gegenphasige Belegung (27) in Gl. (9) ein, so erhält man die Richtcharakteristik des Phasenspiegels

$$|E(\gamma)| = A \cdot \frac{\sin^2\left(\frac{\pi}{2} b \gamma\right)}{\frac{\pi}{2} \gamma}. \quad (36)$$

Bei dem Phasenschieber läßt sich der Ausfall eines schmalen Streifens nahe an $x=0$, sagen wir für $-\frac{a}{2} < x < +\frac{a}{2}$, nicht ganz vermeiden; für diesen Fall ergibt sich, zugleich unter weiterer Verallgemeinerung auf beliebige Phasenverschiebung φ zwischen den beiden Spalthälften als Richtcharakteristik

$$|E(\gamma)| = A \cdot \frac{\sin\left(\frac{\pi}{2} (b-a) \gamma\right)}{\frac{\pi}{2} \gamma} \cdot \cos\left(\frac{\pi}{2} b \gamma \cdot \left(1 + \frac{2a}{b-a}\right) + \frac{\varphi}{2}\right). \quad (37)$$

Gl. (36) ist ein Spezialfall hiervon für $a=0$, $\varphi=\pi$; setzen wir dagegen $\varphi=0$, $a=0$, so folgt die bekannte

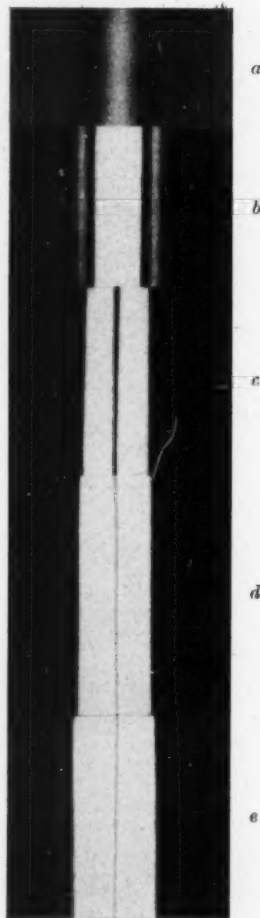


Abb. 2. Aufnahmen zur Minimumstrahlkennzeichnung durch das Phasenverfahren mit Abbildung; a und b Energiestrahle, a kurz, b lang belichtet, c Minimumstrahlkennzeichnung - visueller Eindruck, d Kennzeichnung mit Phasenschieber, e Kennzeichnung mit Phasenspiegel

Formel für den einfachen Spalt

$$|E(\gamma)| = A \cdot \frac{\sin \pi b \gamma}{\pi \gamma} \quad (38)$$

Abb. 3 zeigt diese Richtcharakteristiken in logarithmischem Maßstab, 3a die des gewöhnlichen Spaltes nach Gl. (38), Abb. 3b die des Phasenspiegels nach Gl. (36)

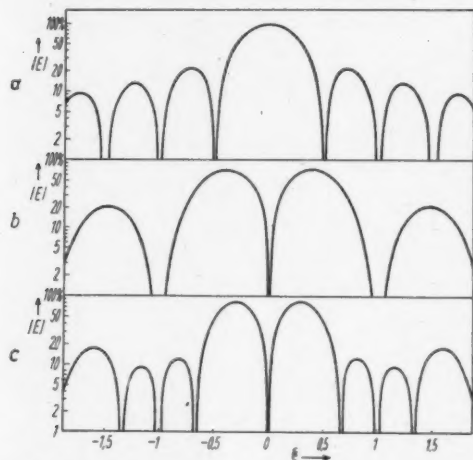


Abb. 3. Richtdiagramme zur Strahlkennzeichnung (theoretisch): *a* bei leerer Aperturbblende (üblicher Energiestrah), *b* mit Phasenspiegel, *c* mit Phasenschieber bei ausfallendem Mittelstreifen

wertung entspricht, wie oben in § 3 gezeigt wurde. Insbesondere können Graubreiten bei Kenntnis der photographischen Weichheit *w* in den verschiedenen Teilen des logarithmischen Richtdiagramms ganz unmittelbar abgelesen werden.

Im interessantesten Bereich der Kurven, der Umgebung der Minima, ist das freilich (bei dem hier benutzten Maßstab) nicht hinreichend genau möglich. Insbesondere ist dort ein Vergleich der Kurven *b* und *c* nicht durchzuführen; dazu zieht man besser den Wert für die Flankensteilheit *s* heran, der sich in beiden Fällen leicht aus dem Momentintegral 1. Ordnung berechnen läßt; man erhält für die nach (27) belegte Blende $s = A \frac{\pi}{2} \cdot b^2$, wie mit Gl. (29) schon ausgesprochen wurde, und bei der ebenfalls gegenphasig belegten Blende mit ausfallendem Mittelstreifen der Breite *a*

$$s = A \cdot \frac{\pi}{2} \cdot (b^2 - a^2). \quad (39)$$

Solange $a < b$, ist das Minimum nur wenig flacher als bei der ungestörten Belegung (27); dieser Einfluß erklärt den Unterschied zwischen Phasenplatte und Phasenspiegel (Abb. 2d und 2e) nur zum kleinen Teil, da *a* nur etwa $\frac{1}{2}$ von *b*, die Abnahme der Flankensteilheit hierdurch also nur 2% ausmachte; daß Mängel der Planparallelplatten des Phasenschiebers den größten Teil dieses Unterschiedes erklären, bemerkten wir schon.

Nimmt *a* aber größere Werte an, nähern wir uns also mit der Blendenbelegung etwa der eines Doppelspaltes, so werden die Flankensteilheiten ersichtlich wesentlich ungünstiger. Gewiß könnte das durch Vergrößerung von *A* kompensiert werden. *A* ist aber durch die Flächenhelligkeit der Lichtquelle und durch die Kohärenztoleranzen (Breite

und Abb. 3c die des Phasenschiebers nach Gl. (37) für $\delta = 180^\circ$ und $a = 0,13 \cdot b$. *A* wurde dabei gleich der Reziproken von $b - a$ gesetzt, damit die Feldstärke eine reine Funktion einer einzigen unabhängigen Variablen

$$\xi = \frac{b - a}{2} \cdot \gamma$$

wurde. Wegen des logarithmischen Maßstabes bedeutet das nur eine Parallelverschiebung gegenüber dem Fall eines für alle drei Kurven gleichen *A*, d. h. gleicher einfallender Intensität.

Der logarithmische Maßstab wurde gewählt, weil er besser als der lineare der tatsächlichen Auswirkung auf das Auge und den Möglichkeiten der photographischen Aus-

des Lichtquellenspaltes) weitgehend beschränkt. Wir wollen hier durchweg annehmen, daß in diesen beiden Punkten bereits alle Möglichkeiten ausgeschöpft sind und können daher aus (39) den großen Vorteil des Phasenverfahrens ($a = 0$) gegenüber dem Doppelspaltverfahren ($a \approx b$) ersehen.

§ 7. Die Verfolgung einer Minimalebene auf einem größeren Teile des Lichtweges; Phasenverfahren und Verdoppelungsverfahren

Wird die Beobachtungsebene E in einer Versuchsanordnung mit Phasenplatte nach Abb. 1 aus der Ebene scharfer Abbildung verschoben, so verbreitern sich die das Minimum flankierenden Maxima; das Minimum selbst bleibt aber — wie der Versuch zeigt und wie aus Symmetriegründen auch ohne weiteres evident ist — einwandfrei. Die Flankensteilheit nimmt dabei ab, aber in geringerem Maße als der Verbreiterung der Maxima entspricht; denn diese nehmen eine Form an, wie man sie von der Fresnelschen Beugung auch sonst gewöhnt ist. Nach einem steilen Anstieg vom Minimum aus nähert sich die Intensität (und zwar bei großem Abstand von der Einstellebene oszillatorisch) einem Grenzwert und fällt dann weiter außen nach den von der Schirmkante her bekannten Oszillationen wieder auf Null ab.

Der mittlere Teil dieses Intensitätsverlaufs ist in Abb. 4d angedeutet. Die in dieser Abbildung skizzierte Versuchsanordnung

enthält die Phasenplatte unmittelbar hinter einer Sammellinse, in deren Brennpunkt der Lichtquellenspalt steht. Das entspricht also dem im letzten Paragraphen behandelten Falle nur dann, wenn die Beobachtungsebene in unendlich großem Abstand von der Phasenplatte liegt. Nähert man die Beobachtungsebene von großen Abständen her der Phasenplatte, so erhöht das die Minimumflankensteilheit auf der Beobachtungsebene $\frac{dE}{d\xi}$, wenn ξ die Koordinate in der Beobachtungsebene ist, in geringerem Maße, als wenn wieder erneut scharf eingestellt würde.

Eine Aufnahme mit 4,13fach vergrößernder Mikrokamera und 5,7facher Nachvergrößerung bei den in Abb. 4 angegebenen Abmessungen der Apparatur zeigt Abb. 5b. Die schwachen Oszillationen sind bei der nach dem Minimum bemessenen Belichtungszeit ($1/2$ Min.) nicht mehr sichtbar. Die zusätzliche probeweise Anbringung einer symmetrischen Spaltblende an der Phasenplatte begrenzte zwar stets die Breite des gesamten Lichtbandes, verbreitert aber das Minimum noch nicht merklich, solange diese Blendenöffnung eine Breite von 1 mm oder mehr

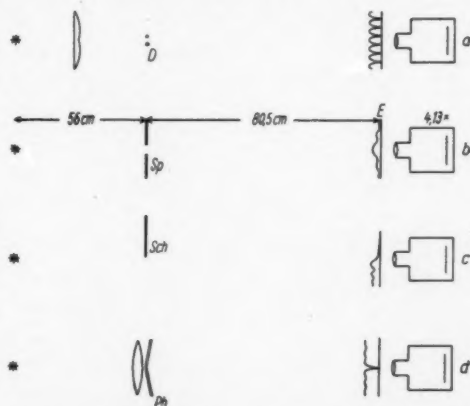


Abb. 4. Versuchsanordnungen zur Strahlkennzeichnung auf einem großen Teile des Lichtweges: a mit reell abbildendem Biprisma, b mit Spalt, c mit Schirmkante, d mit Phasenschieber

hatte. Das ist wesentlich bei der Beurteilung im Vergleich zu den unten zu besprechenden anderen Strahlkennzeichnungen mit einer Aperturblende von 1 mm Breite.

Diese und die folgenden Aufnahmen wurden auf dem Film Isopan *F* bei Rotfilterlicht (Schott RG 1 vor Glühlampe) hergestellt. Der Film arbeitet sehr

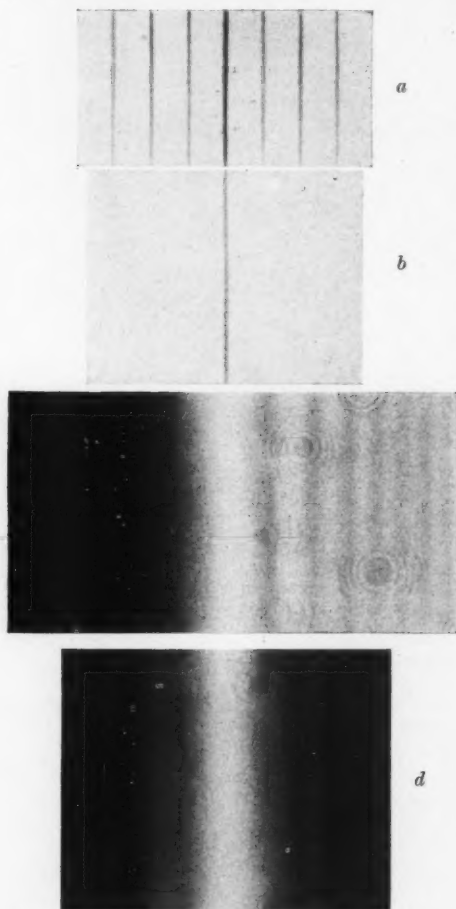


Abb. 5. Aufnahmen zur Strahlkennzeichnung ohne Abbildung, *a* nach dem Verdoppelungsverfahren, *b* nach dem Phasenverfahren mit Phasenschieber, *c* mittels Schirmkante, *d* mit optimalem Spalt

oder einen breiteren Spalt (Abb. 6e) in das Licht einer punktförmigen Lichtquelle stellt und in einem endlichen Abstand auf einem Schirm beobachtet; ξ ist die

weich und zeigt reichlichen Lichthof; der bessere IG-Stat war jedoch nicht mehr beschaffbar. Trotzdem zeigt die Aufnahme einen Gewinn von einer Größenordnung gegenüber der Strahlkennzeichnung durch einen Spalt (Anordnung Abb. 4b, Aufnahme Abb. 5d) oder durch eine Schirmkante (Anordnung Abb. 4c, Aufnahme 5c). Bei dem Vergleich ist noch zu beachten, daß dem Spaltverfahren optimale Bedingungen geschaffen wurden. Die Spaltbreite wurde so eingeregelt, daß die Beugungsfigur in der Beobachtungsebene möglichst scharf wurde; ein breiterer Spalt führte zu einem (von dieser Fresnelschen Beugung her geläufigen) Verschmelzen des Mittelmaximums mit den dann verstärkten Nebenmaxima, ein schmalere Spalt gab zwar Abnahme der Nebenmaxima, aber ein merklich verbreitertes Mittelmaximum. Wurde die Beobachtungsebene *E* bei sonst fest bleibender Anordnung nach rechts oder links verschoben, so wurde der Unterschied zwischen Phasenverfahren und Maximumverfahren noch krasser.

Zum Vergleich seien die theoretischen Richtdiagramme beigelegt (Abb. 6), die man erhält, wenn man ohne jede Abbildung die Phasenplatte (6a), bzw. eine Schirmkante (Abb. 6b), den optimalen Spalt (6c), einen engeren Spalt (6d)

Koordinate auf dem Schirm. Die Überlegenheit der Minimumstrahlkennzeichnung auch für das Fehlen jeder Abbildung ist an den theoretischen Kurven ebenso evident wie an den zugehörigen Aufnahmen.

Muß in irgendeiner Anwendung auf die Abbildung verzichtet werden, wird also dann bei dem Phasenverfahren ohnehin ein Teil der Leistung verschenkt, so werden damit zugleich Verfahren wieder konkurrenzfähig, die dem korrekten Phasenverfahren zwar unterlegen sind, aber den Vorteil besonderer Einfachheit haben und daher mathematisch streng zu erfassen sind. Am vollkommensten ist das der Fall bei der Herstellung der Interferenzminima mittels zweier kohärenter ebener Wellen (oder Kugelwellen oder Zylinderwellen), wie man sie mit einem Fresnelschen Biprisma (oder unbequemer und schlechter mit dem Fresnelschen Doppelspiegel) und einer Linse herstellen kann. Setzt man den Lichtquellenspalt in die Brennebene der Linse, so erhält man äquidistante parallele Nullebenen der Intensität. Oft ist ein System divergenter Nullebenen noch erwünschter, da es z. B. die Untersuchung bei verschiedenen Einfallswinkeln in einer Aufnahme gestattet. Es wird zweckmäßig mit einem reell abbildenden Biprisma (ähnlich den bekannten Billetschen Halbblinsen) realisiert.

Außer Biprismen sind auch andere Anordnungen verwendbar, die aus einer Lichtquelle zwei kohärente Lichtquellen machen. Das Verfahren sei deshalb allgemein als Verdoppelungsverfahren bezeichnet. Prinzipiell scheint ein Doppelspalt auszureichen; praktisch läßt er sich allerdings schlecht mit den erforderlichen Toleranzen herstellen; er besitzt auch energetisch erhebliche Nachteile.

Abb. 5a zeigt eine Aufnahme bei einer Anordnung nach Abb. 4a. Die Unschärfe der Minima ist auch hier um etwa eine Größenordnung kleiner als bei dem Maximum, das mit dem optimalen Spalt als strahldefinierendem Gebilde erzeugt wurde. Wesentlich ist, daß die Spaltbreite bei dem Spaltverfahren und die Basis bei dem Biprismenverfahren (Abstand beider reeller Spaltbilder) gleich waren (1 mm). Selbstverständlich kann man mit größerer Basis bei dem Biprisma viel höhere Genauigkeiten erreichen; der Vergleich erfordert jedoch Verwendung gleicher Aperturblenden.

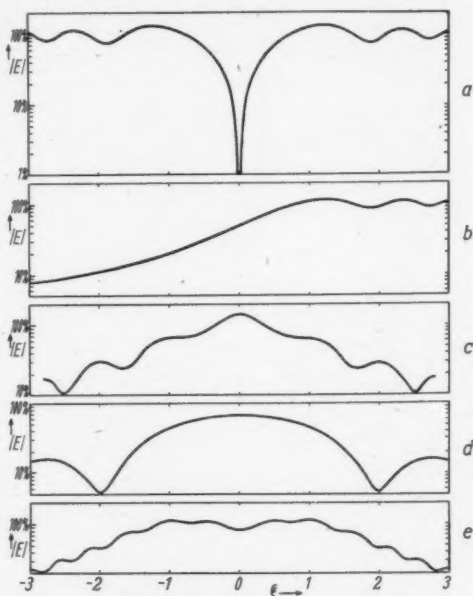


Abb. 6. Theoretische Richtdiagramme zur Strahlkennzeichnung auf einem größeren Teile des Lichtweges, hervorgerufen durch a Phasenplatte, b Schirmkante, c Spalt optimaler Breite, d Spalt von 40% der optimalen Breite, e Spalt von 166% der optimalen Breite

§ 8. Grenzen der Meßgenauigkeit, Einfluß des Störlichtes und der Belichtungszeit

Die mit photographischen Verfahren in verschiedenen Anwendungen erreichten Meßgenauigkeiten entsprechen dem Gewinn von etwa 10 bis 26. (Näheres hierüber folgt in den Einzelarbeiten.)

Das visuelle Verfahren läßt keine volle Ausnutzung der Minimumschärfe zu, da der Schwellenwert des Auges nur einen Eindruck nach Abb. 2c gestattet. Im Gegensatz zur Photographie bringt auch eine Erhöhung der Lichtstärke keine entscheidende Abhilfe, da die hellen benachbarten Maxima das Auge blenden, also sowohl die Pupillenweite verringern als auch den Schwellenwert der Netzhaut anheben. Zweckmäßig verwendet man statt eines Spinnwebfadens in der Gesichtsebene eine V-förmige Blende; sie deckt die hellen Maxima weitgehend ab und läßt eine Unsymmetrie des Minimums innerhalb der Blende besonders in deren schmalsten Stellen deutlich hervortreten.

Das Aussehen der Aufnahmen 2 und 5 läßt ganz unmittelbar vermuten, daß eine weitere Steigerung der Belichtungszeit praktisch keine höhere Minimumschärfe bringt, wie das nach der einfachen Theorie des § 6 zu erwarten wäre; der Versuch bestätigt das. Die Ursache ist offenbar eine Ausfüllung des Minimums durch Streulicht — dazu rechnen wir auch den Lichthof —, durch Kohärenztoleranzen infolge endlicher Breite des Lichtquellenspaltes und Ähnliches. Jedes lineare Minimum verwandelt sich dadurch in ein quadratisches; unsere Theorie bedarf daher einer Ergänzung.

Ein lineares Minimum des Richtdiagramms $|E(\gamma)| = |s \cdot \gamma|$ wird durch eine näherungsweise als ortonunabhängig und in ihrer Auswirkung als inkohärent anzusehende Störfeldstärke E_s in folgender Weise modifiziert

$$|E(\gamma)| = \sqrt{E_s^2 + s^2 \cdot \gamma^2} \approx E_s \cdot \left(1 + \frac{1}{2} \cdot \frac{s^2}{E_s^2}\right); \quad (40)$$

die relative Krümmung K und die absolute Krümmung κ sind also

$$K = \frac{s^2}{E_s^2}; \quad \kappa = \frac{s^2}{E_s}. \quad (41)$$

Während in jedem anderen Extremum (Maximum oder quadratischen Minimum) die absolute Krümmung durch den von der Aperturblende ausgeübten Zwang nach Gl. (16)

$$\kappa \leq \frac{\pi^2}{3} \cdot b^3 \cdot A$$

beschränkt ist, kann am linearen Minimum die absolute Krümmung unbeschränkt ansteigen, wenn nur das Störlicht hinreichend herabgesetzt werden kann. Dadurch wird die Überlegenheit der linearen Minima evident.

Die Graubreite des Minimums erhält man entweder nach Gl. (7) oder ähnlich — aber bei nicht kleinem w etwas genauer — ohne Reihenentwicklung unmittelbar aus Gl. (40)

$$\begin{aligned} (\gamma_h - \gamma_d)_s &= \pm \sqrt{2 \cdot \epsilon^w \cdot \sin w} \cdot \frac{E_s}{s} = \pm \sqrt{2 \cdot \sin w} \cdot \frac{E_g}{s} \\ &= \pm \sqrt{2 \cdot \sin w} \cdot \frac{1}{s} \cdot \sqrt{\frac{G}{C}} \cdot T_s^{\frac{w}{2}}; \end{aligned} \quad (42)$$

darin ist T_s die Belichtungszeit, die dem Störlicht im Minimum die Schwellenbelichtung zuweist. Belichtet man kürzer, so erhält man nach Gl. (40) unmittelbar die Graubreite für beliebiges E_g zu

$$\gamma_h - \gamma_d = \frac{E_g}{s} \cdot \left\{ \sqrt{\epsilon^w - \left(\frac{E_s}{E_g}\right)^2} - \sqrt{\epsilon^w - \left(\frac{E_s}{E_g}\right)^2} \right\}. \quad (43)$$

Gl. (42) ist darin als Spezialfall enthalten für $E_g^2 = E_s^2 \cdot \epsilon^w$, d. h. für Erzeugung der Schwellenbelichtung durch die Störfeldstärke E_s . Setzen wir in Gl. (43) für die

zeit
hten
über

a der
z zur
hilfe,
ver-
man
deckt
rhalb

eine
ringt,
das.
hnen
icht-
n ein

eine
unzu-

(40)

(41)

num)
wang

ränkt
durch

sch —
elbar

(42)

ellen-
elbar

(43)

g der
ir die

Flanke
wir fer
es gil

mit

T ist d
Maxim
Positiv
das Ver
intensit
Minim
nung
Maxim
mumst
kann a
mentell
hältnis
ität z
beiden
kierend
Gl. (45)

g ist
Variabl
gegen
Funktio
blen I_s
über d
Form
mit de
gestellt
konstan
lichtint
stanter
jede K

⁴⁾ V
dieser S
T. Haa
ihm be
nungen
vieler
schickte
geradez
war.

Ann.

Flankensteilheit s den Wert an unserem Minimum $s = A \cdot \frac{\pi}{2} \cdot b^2$ ein und berechnen wir ferner den Gewinn gegenüber dem Maximumverfahren nach Gl. (18) ($m = 1$; es gilt dort dann das Gleichheitszeichen), so erhalten wir

$$g = \frac{\sqrt{\frac{3}{2} \cdot \tau \cdot (1 - e^{-w})}}{\sqrt{1 - I_s \cdot \tau - \sqrt{e^{-2w} - I_s \cdot \tau}}} \quad (44)$$

mit

$$\tau = \left(\frac{T}{t}\right)^p; \quad I_s = \left\{ \frac{E_s}{E_{\text{Spalt Max.}}} \right\}^2 = 1,9 \cdot \left\{ \frac{E_s}{E_{\text{Phas. Max.}}} \right\}^2; \quad (45)$$

T ist darin die für das Minimumverfahren gewählte Belichtungszeit, t die bei dem Maximumverfahren erforderliche, wenn das Maximum gerade weiß auf dem Positiv abgebildet werden soll. I_s bezeichnet die relative Störlichtintensität, d. h. das Verhältnis der Störlichtintensität im Minimum der

Minimumstrahlkennzeichnung zur Intensität des Maximums bei der Maximumstrahlkennzeichnung; es kann auch aus dem experimentell zugänglicheren Verhältnis der Störlichtintensität zur Intensität in den beiden das Minimum flankierenden Maxima nach Gl. (45) entnommen werden.

g ist eine Funktion dreier Variablen w , I_s und τ ; dagegen ist $g \cdot \sqrt{I_s}$ eine reine Funktion der beiden Variablen $I_s \cdot \tau$ und w , kann also über der Abszisse $I_s \cdot \tau$ in Form einer Kurvenschar⁴⁾ mit dem Parameter w dargestellt werden (Abb. 7). Bei konstanter relativer Störlichtintensität I_s und konstanter Weichheit w gibt jede Kurve — bis auf je

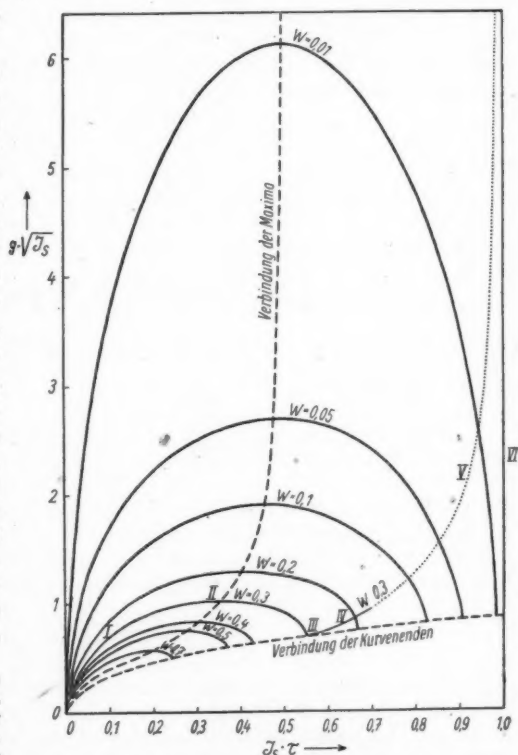


Abb. 7. Durch Minimumstrahlkennzeichnung erreichbarer Gewinn an Graubreitenschärfe g in Abhängigkeit von dem Belichtungszeitverhältnis τ , der Störintensität I_s und der photographischen Weichheit w

⁴⁾ Verfasser möchte an dieser Stelle Herrn stud. phil. T. Haack herzlich danken, der ihm bei numerischen Rechnungen und bei der Anfertigung vieler Zeichnungen ein geschickter, zuverlässiger und geradezu unermüdlicher Helfer war.

einen konstanten Faktor in der Abszisse und der Ordinate — die Abhängigkeit des Gewinns von der Belichtungszeit.

Die Kurve endet rechts, wenn die Belichtungszeit so lang geworden ist, daß die Störlichtintensität den Schwellenwert erreicht; diese Zeit hatten wir oben in Gl. (42) mit T_s bezeichnet, τ_s sei das nach Gl. (45) zugeordnete Zeitverhältnis. Verbinden wir alle rechten Endpunkte der Kurven miteinander, so erhalten wir die gestrichelt eingezeichnete Kurve

$$g_s \cdot \sqrt{I_s} = \sqrt{3} \cdot \frac{(I_s \cdot \tau_s)^{\frac{3}{2}}}{1 + (I_s \cdot \tau_s)^{\frac{3}{2}}}; \quad I_s \cdot \tau_s = e^{-2w} \quad (46)$$

$$g_s = \frac{1}{\sqrt{I_s}} \cdot \frac{1}{2} \cdot \sqrt{3} \cdot \frac{e^{-\frac{3}{4}w}}{\sqrt{\mathfrak{G} \mathfrak{O} \mathfrak{I} \frac{w}{2}}} = \sqrt{\tau_s} \cdot \frac{1}{2} \cdot \sqrt{3} \cdot \frac{e^{\frac{w}{4}}}{\sqrt{\mathfrak{G} \mathfrak{O} \mathfrak{I} \frac{w}{2}}}; \quad \tau_s = \left(\frac{T_s}{t}\right)^p. \quad (47)$$

Da bei kleinem w der Zahlenfaktor hierin rund 1 ist, erfordert ein Gewinn 15 also eine 200fache Belichtungszeit gegenüber dem Energiestrahilverfahren. Abb. 7 zeigt aber auch, daß kürzere Belichtungszeiten nicht nur tragbar, sondern bis zu einem gewissen Grade günstiger sind. Das Maximum der Kurven gibt das Optimum; nach Gl. (44) liegt es vor, wenn

$$E_{g0} = E_s \cdot \sqrt{2 \cdot \mathfrak{G} \mathfrak{O} \mathfrak{I} w}, \quad \text{d. h.} \quad \left(\frac{T_0}{T_s}\right)^p = \frac{e^w}{2 \cdot \mathfrak{G} \mathfrak{O} \mathfrak{I} w} \quad (48)$$

und

$$I_s \cdot \tau_0 = \frac{1}{1 + e^{2w}} \quad \text{mit} \quad \tau_0 = \left(\frac{T_0}{t}\right)^p \quad (49)$$

ist. Die Graubreite ist dann

$$\begin{aligned} (\gamma_h - \gamma_d)_0 &= 2 \cdot \mathfrak{S} \mathfrak{in} w \cdot \frac{E_s}{s} = \sqrt{2 \cdot \mathfrak{S} \mathfrak{in} w \cdot \mathfrak{I} g w} \cdot \frac{E_s}{s} \\ &= \sqrt{2 \cdot \mathfrak{S} \mathfrak{in} w \cdot \mathfrak{I} g w} \cdot \frac{1}{s} \cdot \sqrt{\frac{G}{C} \cdot T_0^{\frac{2}{p}}}; \end{aligned} \quad (50)$$

und der Gewinn ist

$$g_0 = \frac{1}{\sqrt{I_s}} \cdot \frac{1}{2} \cdot \sqrt{\frac{3}{2}} \cdot \frac{\sqrt{1 - e^{-w}}}{\mathfrak{S} \mathfrak{in} w} = \sqrt{\tau_0} \cdot \frac{1}{2} \cdot \sqrt{3} \cdot \frac{e^{\frac{w}{4}}}{\sqrt{\mathfrak{I} g w \cdot \mathfrak{G} \mathfrak{O} \mathfrak{I} \frac{w}{2}}}. \quad (51)$$

Für die Praxis ergibt sich eine sehr einfache Regel zum Auffinden des Optimums: „Durch Belichtungsproben ermittle man die Belichtungszeit; die im Minimum den Schwellenwert der Belichtung ergibt; die optimale Belichtungszeit ist dann rund halb so groß“. Für kurze Belichtungszeiten, bei denen $I_s \cdot \tau \ll e^{-2w}$ ist, besagt (48)

$$g = \sqrt{\tau} \cdot \sqrt{\frac{3}{2}} \cdot \frac{1}{\sqrt{1 - e^{-w}}} \quad (52)$$

in Übereinstimmung mit Gl. (32). Diese für Unterbelichtungen gültige Beziehung hat große praktische Bedeutung, da wir oft gezwungen sind (z. B. bei zeitlich veränderlichen Vorgängen) kürzer zu belichten, als die optimale Belichtung erfordert. Es ist allerdings zu beachten, daß bei sehr kurzer Belichtung die beiden das Minimum umfassenden Graubereiche weiter auseinanderücken, von inhomogenem Streulicht also u. U. verschieden beeinflußt werden und deshalb nicht mehr die volle Ausnutzung der Meßgenauigkeit zulassen (s. unten!).

Über
optimalen
Wird lä
Minimu
bildet. I
Abstand
Anschlu
aus dies
tragen v
fast line
Für die
das Min
abhebt u
tung, ob
genauigh
getragen
Man

I)
II)
III)
IV)
V)
VI)

Die
belichtu
sitätsver
mit einer
sind bei
zeiten de
entweder
anderen
sich mit
scheinen
Die A
Unständ
können.

Für
nicht me
bei klein
den im A
verfahren
Anstieg
nimmt n
 $g \cdot \sqrt{I_s}$
Aufnahm
bei dem
und dem
0,05 hera
der Weic
um rund
werden i
zeichnung
sich eine
filmen er
schichten
lichtung
Die
Kurven
weise bet

Überbelichtungen — so nennen wir alle Belichtungszeiten, die größer sind als die optimale — sind bisher nur so weit erfaßt worden, wie die Kurven in Abb. 7 reichen. Wird länger belichtet, als dem rechten Ende der Kurven entspricht, so liegt die Mitte des Minimums bereits in dem Graubereich und die schwarze Grenze wird nicht mehr abgebildet. In Verallgemeinerung des Begriffs verstehen wir dann unter Graubreite den halben Abstand der beiden das Minimum einschließenden Hellgrenzen; das ist sinngemäß, da stetiger Anschluß an den bisherigen Gebrauch des Begriffes auf diese Weise gesichert ist. Der sich aus dieser Graubereichsdefinition ergebende Gewinn ist nur für $w = 0,3$ in Abb. 7 eingetragen worden (zu viele Kurven würden die Übersicht stören) und zeigt einen zunächst fast linearen und dann ins Unendliche führenden Anstieg; der Pol liegt bei $I_s \cdot \tau = 1$. Für die Praxis hat der stark ansteigende Teil der Kurven keinen rechten Sinn, da sich das Minimum dort nur durch sehr geringe Schwärzungsdifferenzen gegen die Umgebung abhebt und viel weniger genau zu messen ist als bei optimaler oder sogar zu kurzer Belichtung, obgleich die hohen Ordinaten der Kurve einen großen Gewinn und also große Meßgenauigkeit erwarten ließen. Die Kurve wurde in diesem Teile daher nur punktiert eingetragen.

Man unterscheidet zweckmäßig 6 Bereiche der Belichtungszeit:

- I) Unterbelichtung: Stark ansteigender Teil der Kurven links.
- II) Optimale Belichtung: Umgebung des Maximums; $g \geq 0,8 \cdot g_0$.
- III) Schwache Überbelichtung: Stark abfallender Teil der Kurve bis zur Spitze.
- IV) Mäßige Überbelichtung: Erneuter fast linearer Anstieg der Kurve; $g \leq g_0$.
- V) Starke Überbelichtung: Starker bis fast senkrechter Anstieg der Kurve.
- VI) Totbelichtung: Das Gebiet $I_s \cdot \tau \geq 1$, in dem keine Differenzierung des Minimums gegen die Umgebung wahrnehmbar.

Die Gebiete V und VI sind praktisch unbrauchbar. Das Gebiet mäßiger Überbelichtung IV eignet sich zur Veranschaulichung besser als jedes andere, da es den Intensitätsverlauf im Minimum selbst noch erfaßt. Die Aufnahmen zu Abb. 2 wurden daher mit einer in dieses Gebiet fallenden Belichtungszeit hergestellt. Für praktische Messungen sind bei den speziellen Arbeiten mit der Minimumstrahlkennzeichnung stets Belichtungszeiten der Gebiete I, II oder III benutzt worden und zwar des Gebietes I nur dann, wenn entweder die langen Belichtungszeiten des Gebietes II nicht tragbar waren oder die in anderen Teilen der Apparatur liegenden Meßfehlergrenzen eine volle Ausnutzung der an sich mit der Minimumstrahlkennzeichnung erreichbaren Genauigkeit nicht erforderlich scheinen ließen.

Die Abgrenzung der Gebiete enthält natürlich eine gewisse Willkür; so wird man unter Umständen andere Zahlenfaktoren an g_0 bei Definition der Gebiete II und IV anbringen können.

Für größere Weichheit w lassen sich die verschiedenen Gebiete der Belichtungszeit nicht mehr voneinander trennen; die richtige Wahl der Belichtungszeit wird daher erst bei kleiner Weichheit kritisch. Von einer Herabsetzung der Weichheit, wie man sie mit den im Anhang angegebenen Mitteln erreichen kann, hat natürlich auch jedes Maximumverfahren Vorteil, aber ein Minimumverfahren in weit höherem Maße; das zeigt der starke Anstieg des optimalen Gewinns g_0 im Scheitel der Kurven der Abb. 7. Beispielsweise entnimmt man für $w = 0,3$ und $I_s = 1/800$ aus der Abb. 7, daß bei $I_s \cdot \tau = 0,25$ die Ordinate $g \cdot \sqrt{I_s} = 1$, also $g = 24$ ist. Das entspricht den praktischen Verhältnissen, die bei unseren Aufnahmen zu Abb. 2 vorlagen. Die erforderliche Belichtungszeit T ist 75 sec, wenn die bei dem Maximumverfahren benötigte 0,5 sec beträgt. Bei der gleichen Belichtungszeit und demselben Störlichtverhältnis kann man dagegen mit einer Weichheit w , die auf 0,05 herabgedrückt ist, nach Abb. 7 einen Gewinn $g = 58$ erreichen. Diese Verringerung der Weichheit w setzt nach Gl. (18) auch die Graubreite des Maximumverfahrens bereits um rund den Faktor 3 herab; daher kann die Genauigkeit um rund das 170fache gesteigert werden im Vergleich zu einem Verfahren, bei dem man weder die Minimumstrahlkennzeichnung noch die chemischen Kunstgriffe der Filmbehandlung anwendet. Leider läßt sich eine so weitgehende Herabsetzung der Weichheit nur bei sehr feinkörnigen Spezialfilmen erreichen, die lange Belichtungszeiten erfordern, oder bei den üblichen Bromsilberschichten unter entsprechend starker Vorvergrößerung, die ihrerseits ebenfalls lange Belichtung bedingt.

Die vorstehenden Ergebnisse über die zweckmäßigsten Belichtungszeiten und die Kurven der Abb. 9 behalten auch im wesentlichen ihren Sinn, wenn statt des vorzugsweise betrachteten Phasenverfahrens eine andere Minimumstrahlkennzeichnung vorliegt.

Der Unterschied liegt lediglich darin, daß dann der Gewinn g und die Belichtungszeit T nur noch relativen Charakter haben, da das Maximumverfahren als Vergleichsverfahren nicht unmittelbar herangezogen werden kann und gegebenenfalls durch ein anderes ersetzt werden muß. Der Verlauf der Kurven wird dadurch nicht geändert und damit auch nicht die Einteilung der Belichtungszeiten in die Gruppen I bis VI. Für quantitative Betrachtungen geht man dann zweckmäßig auf die unter den hier mitgeteilten Gleichungen zurück, in denen die Flankensteilheit s noch explizit auftritt.

Sind die Ergebnisse also in dieser Hinsicht einer Verallgemeinerung ohne weiteres fähig, so bedürfen sie einer wesentlichen Einschränkung bezüglich der Voraussetzungen über das Störlicht, die in diesem Paragraphen auf S. 356 gemacht wurden. Es ist nötig, folgende Arten von Störlicht zu unterscheiden:

- a) Inkohärent wirkendes Störlicht in annähernd gleichmäßiger Verteilung über große Teile des Gesichtsfeldes,
- b) Kohärentes Streulicht, annähernd gleichmäßig über große Teile des Gesichtsfeldes verteilt,
- c) Unregelmäßig über das Gesichtsfeld verteiltes Streulicht,
- d) Lichthof,
- e) Durch endliche Breite des Lichtquellenspaltes verursachte Inkohärenz der sich in der Beobachtungsebene überlagernden Wellenzüge,
- f) Durch die Breite des Spektralbereichs bedingte gegenseitige Störung von Beugungsbildern verschiedener Wellenlänge.

Licht der Gruppe a und b stammt von Streuzentren aus großen, c aus mittleren Entfernungen der Beobachtungsebene und d aus der photographischen Schicht selbst. Die Voraussetzungen unserer Rechnung treffen streng nur für die Gruppe a zu. Doch wirken sich der Lichthof und die Störungen der Gruppen e und f sehr ähnlich aus; setzt man statt ihrer eine ortsunabhängige und inkohärente Störintensität in der im Minimum vorliegenden Größe voraus, wie das geschehen ist, so hat man damit höchstens die Minimumstrahlkennzeichnung etwas zu ungünstig beurteilt, da die Flankensteilheit im Minimum in der Praxis auf Grund des Lichthofes usw. stets größer ist, als damit angesetzt. Streulicht der Gruppe c , das z. B. in den Aufnahmen 5c die mehrfach zu beobachtenden konzentrischen Kreisinge hervorrief, ist zwar bei Aufnahmen eines Minimums wegen der großen Flankensteilheit meist nicht als solches erkennbar; doch bewirkt es unter Umständen eine leichte Schlingelung der das Minimum anzeigenden Geraden. Bei einfachen Lichtzeigergeräten ließ es sich leicht vermeiden; doch trat diese zusätzliche Maßfehlerquelle deutlich in der Arbeit ^{1a)} hervor, in der das Licht mehr als 50 Flächen durchsetzte, die alle als Träger von Streuzentren in Frage kamen.

Merkliches Streulicht der Gruppe b müßte als besonders gefährlich für die Minimumstrahlkennzeichnung angesehen werden. Während ein Maximum dadurch nur verbreitert würde, muß man bei einem linearen Minimum eine Parallelverschiebung erwarten, da das Licht zu beiden Seiten des Minimums um 180° in Phase gegeneinander versetzt ist und verschieden beeinflusst würde. Das könnte unter Umständen Anlaß zu systematischen Fehlern geben. Inwieweit Streulicht dieser Art auftritt, läßt sich besonders einfach mit dem Verdoppelungsverfahren untersuchen, das statt eines einzelnen Minimums eine große Zahl äquidistanter Minima besitzt, die sich bei Hinzutreten kohärenten Streulichtes gegenläufig zueinander verschieben müßten. Durch Messung der Minimumabstände läßt sich die Wirkung dieses kohärenten Streulichtes abschätzen und nötigenfalls durch Mittelbildung eliminieren. Durch Vermessen der Negative im Komparator konnte auf diese Weise festgestellt werden, daß Streulicht dieser Art innerhalb der durch die Breite der Minima (Unschärfe 1:10 bis 1:20) gesteckten Genauigkeitsgrenzen nicht auftrat.

In den weitaus meisten Fällen sind die Gruppen d , e und f entscheidend. Diese Störungen lassen sich zwar grundsätzlich stets beseitigen; doch würde das zu ganz außerordentlich langen Belichtungszeiten führen. Insbesondere kann der Lichthof durch hinreichende Vorvergrößerung weitgehend unschädlich gemacht werden. Ähnlich kann bei hinreichend großem Aufwand an Belichtungszeit auch die durch die Gruppen e und f verursachte Unschärfe prinzipiell beliebig weit ausgeschaltet werden, da sowohl der Lichtquellenspalt als auch der Spektralbereich sehr eng gemacht werden können. In welchem Umfang das tatsächlich erforderlich ist, soll im nächsten § gesondert betrachtet werden, da dieses Problem in der Literatur nach Regeln geschah, deren Allgemeingültigkeit überschätzt wurde.

§ 9. Kohärenztoleranzen

Strenge Kohärenz der Lichtquelle erfordert einen äußerst schmalen Lichtquellenspalt. Wird dieser auf δx_* verbreitert, so verbreitert sich das den „Strahl“ kennzeichnende Gebilde (Minimum o. a.) in der Beobachtungsebene ebenfalls durch Überlagerung verschiedener Beugungsbilder. Nach dem sehr allgemein geltenden Satze der Optik, der in der Fußnote ³⁾ auf S. 342 zitiert wurde, geschieht diese Verbreiterung um ein Stück δx , das mit den ausgenutzten Aperturen an der Lichtquelle $\delta \gamma_* \cdot \lambda$ und in der Beobachtungsebene $\delta \gamma \cdot \lambda$ durch die Gleichung von der Invarianz der Unschärfe

$$\delta x_* \cdot \delta \gamma_* = \delta x \cdot \delta \gamma$$

verknüpft ist, ganz unabhängig davon, welche idealen abbildenden Mittel noch zwischen Lichtquellenspalt und Beobachtungsebene sich befinden. Läßt nun die Strahlkennzeichnung noch Verschiebungen der Dimension

$$\delta x \cdot \delta \gamma = \frac{1}{g}$$

erkennen, so muß also die Kohärenzbedingung für die Lichtquelle

$$\delta x_* \delta \gamma_* \ll \frac{1}{g} \quad (53)$$

lauten, sofern eben die endliche Breite der Lichtquelle innerhalb der experimentellen Feststellbarkeit nicht störend sein soll.

Zu dieser „allgemeinen lateralen Kohärenzbedingung“ muß also die übliche Verdetsche ⁴⁾ Kohärenzbedingung verschärft werden. Die Verdetsche Bedingung lautete in unseren Bezeichnungen $\delta x_* \delta \gamma_* \ll 1$. Ihre Begründungen, daß die Breite des Spaltes sich im Bilde dann der „experimentellen Feststellbarkeit“ entzöge, waren dabei der Aussage äquivalent, daß die experimentelle Feststellbarkeit durch die Unschärfebedingung $\delta x \delta \gamma \geq 1$ begrenzt sei (in unserer Sprechweise). In demselben Maße, wie die Minimumstrahlkennzeichnung die Unschärfebedingung unterbietet, sind nun auch die Kohärenztoleranzen zu verringern.

Praktisch nötigt Gl. (53) uns also, den Lichtquellenspalt um den Faktor $1/g$ schmäler zu machen, als das bei dem üblichen Energiestrahle nötig wäre. Die oben berechneten Beziehungen zwischen Belichtungszeit und Gewinn setzten gleichbleibenden Lichtquellenspalt voraus. Wählt man die Spaltbreite aber stets so groß, wie gerade bei dem jeweiligen Verfahren noch tragbar ist, so sind jene Formeln in dem Sinne zu modifizieren, daß überall statt $\tau = \left(\frac{T}{t}\right)^p$ die Größe $\tau' = \left(\frac{T}{t}\right)^p \cdot \frac{1}{g}$ einzusetzen ist. Verzichtet man mit Rücksicht auf die Belichtungszeit auf volle Ausnutzung der Strahlschärfe, indem man den Lichtquellenspalt breiter macht, als der Kohärenzbedingung entspricht, so erhält man einen resultierenden Gewinn g_r , der sich nach

$$g_r = \frac{1}{\sqrt{\frac{1}{g^2} + (\delta x_* \cdot \delta \gamma_*)^2}} \quad (54)$$

berechnet.

³⁾ M. Berek, Z. Physik **53**, 483 (1929); K. W. Kohlrausch, Ausgewählte Kapitel aus der Physik, II. Teil, Optik, Wien 1948 und fast alle anderen Lehrbücher der Optik.

Die linke Seite von (53) wollen wir als „laterale Inkohärenz“ der Lichtquelle bezeichnen. (53) besagt dann in sehr anschaulicher Form, daß die Inkohärenz der Lichtquelle klein gegenüber der Unschärfe der benutzten Strahlkennzeichnung sein muß, damit die Inkohärenz sich praktisch nicht auswirkt.

Die Breite des benutzten Spektralbereichs $\delta\lambda$ verbreitert das den Strahl kennzeichnende Gebilde (speziell das Minimum) nicht merklich, wenn

$$\frac{\delta\lambda}{\lambda} \ll \frac{1}{n \cdot g} \quad (55)$$

ist. Darin ist n die „Ordnung“, also der Gangunterschied der miteinander interferierenden beiden Wellenzüge gemessen in Wellenlängen; n ist bei der Phasenplatte $\frac{1}{2}$, bei dem Biprisma $\frac{1}{2}$, $\frac{3}{2}$, usw. je nach Zahl der benutzten Minima. Diese „Laufzeitkohärenzbedingung“ wird aus der bekannten Bedingung dieser Art leicht gewonnen durch die hier ersichtlich wieder erforderliche Verschärfung. Bei einem Gewinn $g = 100$ und Strahlkennzeichnung mit Phasenplatte muß im Sichtbaren also $\delta\lambda \ll 100 \text{ \AA}$ sein. Die grüne Linie einer Quecksilberhochdrucklampe ist also sogar bei sehr viel höheren Ordnungen und zur Erzielung noch höherer Gewinne ausreichend. Bei geringeren Gewinnen genügen oft Farbfilter oder die selektive Sensibilität der blauempfindlichen Filme.

§ 10. Subtilität, Meßbarkeit und Nachweisbarkeit eines Effektes

Neben den soeben eingeführten Begriffen und der in § 2 definierten „Schärfe“ einer Strahlkennzeichnung erweisen sich in den Anwendungen als nützlich quantitative Fassungen der Begriffe Subtilität, Meßbarkeit und Nachweisbarkeit eines Effektes. Sie ermöglichen, Gedankenarbeit durch formales Rechnen zu ersetzen und sonst nur umständlich beschreibbare Zusammenhänge kurz zu formulieren.

Z. B. ist die Parallelversetzung Δx , die ein „Strahl“ bei der Totalreflexion erfährt, von seiner Apertur, sagen wir von seiner Winkelstreuung im Maß der Richtvariablen $\Delta\gamma$, abhängig. Unter den in der Arbeit¹⁾ im Einzelnen diskutierten Verhältnissen ist nun $\Delta x \cdot \Delta\gamma \ll 1$. Dieser Effekt ist also ohne besondere Maßnahmen mit den üblichen Strahlen, wie sie durch Spalte u. ä. realisiert werden, nicht mehr meßbar. Wir nennen ihn deshalb einen subtilen Effekt und bezeichnen quantitativ als seine „Subtilität“ die Größe

$$\sigma = \frac{1}{\Delta x \cdot \Delta\gamma} = \frac{\lambda}{\Delta x \cdot \Delta\alpha}; \quad (56)$$

Allgemein ist ein Effekt der Subtilität σ ohne Vervielfachung nur mit einer Strahlkennzeichnung der Schärfe $g \geq \sigma$ meßbar.

Bezeichnen wir als Meßbarkeit μ eines Effekts mit einer Strahlschärfe $g = \frac{1}{\delta x \cdot \delta\gamma}$ die Größe

$$\mu = \frac{g}{\sigma} = \frac{\Delta x \cdot \Delta\gamma \text{ (Effekt)}}{\delta x \cdot \delta\gamma \text{ (Meßmethode)}}, \quad (57)$$

so gilt also der Satz: Ein Effekt kann mit einer Methode nur gemessen werden, wenn die Meßbarkeit $\mu \geq 1$ ist. $\mu = 1$ bedeutet offenbar, daß Effekt und Strahlschärfe gleich sind; daher kann erfahrungsgemäß bei $\mu = 1$ noch eine Messung auf 10 bis 20% mittlerem Fehler der Einzelmessung als möglich angesehen werden.

Die „Nachweisgrenze“, wie wir in ohne weiteres verständlicher Weise sagen wollen, liegt unter 1; wir definieren sie bei $\mu = 0,2$.

§ 11. Beispiele für die Anwendbarkeit der Minimumstrahlkennzeichnung

Die Minimumstrahlkennzeichnung erwies sich als besonders nützlich bei der Untersuchung des Goos-Hänchen-Effekts in der in Fußnote 1 zitierten Arbeit. Da die seitliche Versetzung, die ein Strahl bei der Totalreflexion erfährt, sehr klein ist, ist die Subtilität groß; selbst in günstigen Fällen betrug sie 35. Goos und Hänchen konnten den Effekt zwar nach 34facher Vervielfachung (durch wiederholte Reflexion) nachweisen; die Abhängigkeit von der Polarisation entging ihnen*), da deren Subtilität noch größer ist. Mit der Minimumstrahlkennzeichnung wird jedoch auch diese Polarisationsabhängigkeit mit noch befriedigender Genauigkeit meßbar, obwohl die Vervielfachung nicht so weit getrieben wurde wie bei Goos und Hänchen.

In ähnlicher Weise kann man der kleinen Spiegeldrehung eines Spiegelgalvanometers eine Subtilität zuordnen, die beispielsweise 25 ist, wenn der zu messende Strom nur noch 4% der mit dem üblichen Lichtzeiger meßbaren Stromstärke beträgt. Dieser Fall wurde oben wegen seiner Übersichtlichkeit und Einfachheit mehrfach als Beispiel herangezogen; tatsächlich interessiert er unter den Anwendungen relativ am wenigsten, da eine ähnliche Wirksamkeit schon von anderen Mitteln (Galvanometerverstärkungen usw.) bekannt ist.*

Am überraschendsten sind die Erfolge der Minimumstrahlkennzeichnung in solchen Fällen, bei denen es sich nicht einfach um eine Parallelverschiebung oder Schwenkung des Strahles handelt, bei denen vielmehr noch eine zusätzliche Transformation stattfindet, so daß ganze Kurven mit der Minimumstrahlkennzeichnung dargestellt werden. In der Arbeit*) wurde über eine Verbesserung in der Vermessung von Schlieren und ähnlichen berichtet, die hierauf beruht. Auch die Ablenkung des Lichtes durch eine schwache Schliere ist als subtiler Effekt anzusehen.

Die Minimumstrahlkennzeichnung mit einer 180°-Phasenplatte läßt in gewissen Fällen eine Verbesserung mikroskopischer Beobachtungen zu. Ferner ist eine Spektrallinie, deren Breite unter dem Auflösungsvermögen eines benutzten Spektographen liegt, für diesen ein subtiles Gebilde. Setzt man in den parallelen Teil des Strahlenganges eine Phasenplatte, so kann die Schärfe des Minimums ein Maß der Spektrallinienbreite liefern. Messungen über die Lage der Linie lassen sich so ebenfalls verbessern.

Welches der beiden Verfahren, Phasenverfahren oder Verdoppelungsverfahren, zur Anwendung kommt, hängt natürlich stets von dem jeweils vorliegenden Problem ab. Gelegentlich sind beide Verfahren verwendbar. So kann für ein Drehspiegelgerät nicht nur das oben dafür diskutierte Phasenverfahren benutzt werden, sondern auch das Verdoppelungsverfahren, z. B. so, daß nach einer der oben beschriebenen Anordnungen Nullebenen über den Spiegel auf die Beobachtungsebene geworfen werden, vorzugsweise derart, daß die reellen Bilder des Lichtquellenspaltes fast an die Ränder des Drehspiegels fallen. Eine ähnliche Einrichtung ist auch bei Schlierenverfahren verwendbar. In anderen Fällen, in denen die Verwechslungsmöglichkeiten des Verdoppelungsverfahrens stören, oder z. B. bei dem Spektrometer ist das Phasenverfahren das angemessene; im letzten Falle ist ja eine Abbildung der erforderlichen Art bereits vorhanden.

*) Gleichzeitig mit dem Verf. berichtete F. Goos jedoch auf der Hamburger Physikertagung von der Beobachtung einer Polarisationsabhängigkeit; eine 70fach wiederholte Reflexion hatte zum Ziel geführt.

§ 12. Ersatz der Begriffe Strahl und Lichtweg durch die Minimumstrahlkennzeichnung. Mathematische Vorteile

Außer der Genauigkeitssteigerung gewinnt man bei manchen optischen Phänomenen eine klare und eindeutige Fragestellung erst durch die Minimumstrahlkennzeichnung.

So ist es z. B. fraglich, was man unter der Strahlversetzung bei der Totalreflexion verstehen will, solange man einen „Energistrah“ verwendet. Denn dieser hat nach der Totalreflexion eine völlig veränderte Intensitätsverteilung in seinem Querschnitt. Die vorher symmetrischen Nebenmaxima sind nach der Reflexion unsymmetrisch, und das hellere der beiden Nebenmaxima 1. Ordnung ist mit dem Hauptmaximum verschmolzen. Man könnte nun z. B. das absolute Maximum auf seinem Wege verfolgen und als „Strahl“ ansehen; oder man könnte den jeweiligen Schwerpunkt der Energie als maßgebend ansehen; schließlich könnte man noch von vornherein bereits ein sehr breites Lichtbündel als „Strahl“ benutzen, die Versetzungen der „Kanten“ (definiert etwa als Orte mit $\frac{1}{4}$ der Intensität) berechnen bzw. messen und das Mittel aus beiden zur Strahlversetzungsdefinition verwenden. Alle diese Definitionen wären verschieden in ihrer Auswirkung, und jede wäre überdies von der speziellen Form des „Ausgangsstrahles“ abhängig. Letzten Endes wäre der Vorgang eben nicht durch die Angabe einer Strahlversetzung hinreichend beschreibbar; vielmehr müßte der Einfluß der Totalreflexion auf das Wellenbündel durch die Intensitätsverteilung im reflektierten Wellenbündel beschrieben werden.

Benutzt man dagegen als Strahlersatz Ebenen verschwindender Intensität (Nullebenen), wie sie durch Interferenz zweier ebener Wellen entstehen (Verdoppelungsverfahren), so geht dieses System durch die Totalreflexion wieder in ein kongruentes System ebensolcher Nullebenen über, die gegenüber den „elementar geometrisch reflektierten“ Ebenen um eine wohldefinierte Strecke verschoben sind. Statt der Intensitätsverteilung im Ausgangsstrahl genügt hier also die Angabe des Kreuzungswinkels, unter dem sich die beiden ebenen Wellen schneiden, und das Ergebnis ist statt durch eine neue Energieverteilung sehr einfach durch Angabe jener Versetzung der Ebenen bereits vollständig zu beschreiben¹⁾.

In anderen Fällen, in denen zwar ein kongruentes Ebenensystem resultiert, die Intensität im Minimum aber von Null verschieden ist nach dem Passieren der reflektierenden (oder allgemeiner das Ebenensystem beeinflussenden) Vorrichtung, ist noch die Angabe einer zweiten Zahl, der Doppelwertsbreite des Minimums, erforderlich. So ist z. B. die metallische Reflexion durch die Strahlversetzung und die Minimumbreite zu charakterisieren.

Ein dritter Vorteil der Minimumstrahlkennzeichnung liegt in der mathematischen Einfachheit. So könnte z. B. der Goos-Hänchen-Effekt für den Energistrah nicht streng berechnet werden²⁾. Artmann vermutete daher irrtümlich die Ursache für die Diskrepanz zwischen seiner Rechnung und den Goosschen Messungen in seiner Näherung (im wesentlichen in dem Kirchhoffschen Ansatz).

Dagegen ließ sich der entsprechende Effekt für die Minima des Verdoppelungsverfahrens völlig streng berechnen, und das Ergebnis befand sich in voller Übereinstimmung mit den Messungen. Das Schicksal, das zwei ebene Wellen beim Passieren auch komplizierter Vorrichtungen erleiden, ist naturgemäß einfacher

²⁾ K. Artmann, Ann. Physik (6) 2, 87 (1948).

mathem
bündels

Ein
man vo
diese ze
mehr v
Gele

wellen,
Wesen
variatio
entsteh
Ergebn
welle) le
graphis
wellen)

Ver
tierende
der Aus
über de
zisieren
minimu

Die
möglich
geben
wir hier

Die
Physik
läufiger
Währen
durch
schriebe
nicht is
die For
definier
Minimu

Die
bei To
„Ferna
Orten
der W
Photon
Nullins
äußerst

Auf
langer
Übertra

mathematisch zu erfassen, als das eines bereits kompliziert aufgebauten Wellenbündels, wie ein Energiestrahle es ist.

Eine noch weitere Vereinfachung ist grundsätzlich nicht mehr denkbar, da man von zwei ebenen Wellen dann schon zu einer ebenen Welle übergehen müßte; diese zeichnet aber selbstverständlich keine Ebene, keine Gerade und keinen Punkt mehr vor einem anderen aus.

Gelegentlich kann die Verwendung zweier anderer Elementarwellen (Kugelwellen, Zylinderwellen) statt ebener Wellen dem Problem angemessener sein. Das Wesen der Minimumstrahlkennzeichnung bleibt dabei stets eine kleine Ortsvariation des Quellgebietes und Differenzbildung der vor und nach der Variation entstehenden Amplituden. Im Limes kleiner Variation kann man daher stets das Ergebnis der Minimumstrahlkennzeichnung (aus dem Ergebnis für eine Elementarwelle) leicht durch Differentiation gewinnen. Die Rechnung läßt sich oft auch durch graphische Verfahren (Schnittbildung zwischen Phasenflächen zweier Elementarwellen) weitgehend ersetzen.

Verfasser vermutet, daß so z. B. die Frage nach der Zuordnung der resultierenden Wellen (Kopfwelle usw.) bei Reflexion einer Kugelwelle zu Teilgebieten der Ausgangswelle einen Sinn bekommen kann. Ebenso lassen sich manche Fragen über den Weg elektromagnetischer Wellen bei Reflexion an der Ionosphäre präzisieren und mathematisch (auch experimentell) vereinfachen, wenn man mit minimumgekennzeichneten Wellen arbeitet.

§ 13. Analogien zu anderen physikalischen Gebieten

Die Minimumstrahlkennzeichnung wurde hier unmittelbar aus der Forderung möglichst scharfer Kennzeichnung einer Ebene im Raum begründet. Vielleicht geben aber einige Analogiebetrachtungen zu anderen Gebieten der Physik, die wir hier zum Schluß skizzieren wollen, einen anschaulicheren Zugang.

Die Einführung des folgenden Vergleichs scheint berechtigt, da heute dem Physiker die Folgerungen aus der Heisenbergschen Unschärfebedingung geläufiger sind als die aus deren klassischem Vorläufer, der Unschärfebedingung (1). Während z. B. im Wasserstoffatom der Aufenthaltsort des Elektrons stets nur durch eine Wahrscheinlichkeitsverteilung über ein dreidimensionales Gebiet beschrieben werden kann, bilden die Orte, an denen das Elektron mit Sicherheit nicht ist, nur eine zweidimensionale Mannigfaltigkeit; d. h. man kann Flächen durch die Forderung einer Aufenthaltswahrscheinlichkeit Null mit unbegrenzter Schärfe definieren. Dem entspricht das Verhältnis der Energiestrahledefinition zu der Minimumstrahlkennzeichnung.

Die Minimumstrahlkennzeichnung ist eine Nullmethode. Während man z. B. bei Toeplers Schlierenverfahren die Orte sucht, in denen die Intensität $\frac{1}{4}$ der „Fernabintensität“ ist, fragt man bei der Minimumstrahlkennzeichnung nach den Orten der Intensität Null. Das macht wie bei anderen Nullmethoden (z. B. bei der Wheatstoneschen Brücke) ein quantitatives Nachweismittel (hier die Photometrie) unnötig. Der Verwendbarkeit eines hochempfindlichen ungeeichten Nullinstruments bei elektrischen Nullmethoden entspricht hier die Möglichkeit äußerster Überbelichtung, die das Minimum mit großer Schärfe erkennen läßt.

Auf anderen Gebieten der Physik und Technik sind Minimummethoden seit langer Zeit im Gebrauch. So kann die Minimumstrahlkennzeichnung auch als eine Übertragung der Minimumpeilverfahren der Hochfrequenztechnik auf die Optik

betrachtet werden. Es ist aus der Unschärfebedingung (1) evident, daß die Minimumstrahlkennzeichnung um so nötiger erscheint, je länger die benutzten Wellen sind.

§ 14. Anhang

A) Neben den üblichen Begriffen wie Schwellenwert und Kontrastfaktor leisten bei photographischen Arbeiten einige andere Begriffe gute Dienste, die auch in dieser Arbeit benutzt wurden und deren Zusammenhang mit den üblichen Begriffen hier hergestellt werde.

1. Die Eigenschaften eines Negativmaterials, soweit sie hier in Frage kommen, werden zweckmäßig erfaßt durch: den Schwellenwert, der hier als die Mindestbelichtung B_d definiert sei, die noch eine merkliche Schwärzung hervorruft, und den Kontrastfaktor Γ_n , der gleich dem Differentialquotienten der Schwärzung nach dem dekadischen Logarithmus der Belichtung im unteren — geradlinig ausgemittelten — Teil der Schwärzungskurve sei. Der Schwellenwert weicht dabei also von der Inertia nach Hurter und Driffield ab; der Kontrastfaktor ist meist kleiner anzusetzen als der entsprechende Wert nach Hurter und Driffield, da der hier wichtige untere Teil der Schwärzungskurve eine im Mittel kleinere Steigung hat als der geradlinige Teil.

2. Als dekadische Weichheit eines Positivmaterials v_p bezeichnen wir den dekadischen Logarithmus des Opazitätsverhältnisses (d. h. den Schwärzungsunterschied) zweier Negativstellen, von denen die eine auf dem Positivmaterial gerade noch „weiß“, die andere gerade „schwarz“ abgebildet wird.

3. Die Kopie oder Vergrößerung eines Negativs auf einem Positivmaterial werde so vorgenommen, daß der Schwellenwert des Negativs gerade aus dem tiefen Schwarz sichtbar wird.

Dann werde als dekadische Weichheit v des kombinierten Negativ-Positiv-Verfahrens der dekadische Logarithmus des Verhältnisses zweier Negativbelichtungen bezeichnet, von denen die eine gerade aus dem tiefsten Schwarz des Positivs und die andere gerade aus dem reinen Weiß sich merklich abheben. Die erste dieser Belichtungen ist der Schwellenwert B_d , die andere heiße die Vollbelichtung B_h . Als Graubelichtung G des Gesamtverfahrens bezeichnen wir das geometrische Mittel aus Vollbelichtung und Schwellenwert.

Zwischen diesen Größen bestehen dann offenbar folgende Beziehungen

$$B_h = B_d \cdot 10^v; \quad G = B_d \cdot 10^{\frac{v}{2}}; \quad v = \frac{v_p}{\Gamma_n}. \quad (65)$$

Außer der dekadischen Weichheit sei noch eine Weichheit schlechthin definiert durch

$$w = \frac{v}{2} \cdot \ln 10 = v \cdot 1,15; \quad w_p = \frac{v_p}{2} \ln 10 = v_p \cdot 1,15; \quad (66)$$

dann gilt

$$w = \frac{w_p}{\Gamma_n}; \quad B_d = G \cdot e^{-w}; \quad B_h = G \cdot e^w. \quad (67)$$

Für die Zwecke dieser Arbeit genügt es, die Weichheit des kombinierten Negativ-Positiv-Verfahrens zu messen; das kann auch ohne Graukeil, Sensitometer usw. geschehen, indem z. B. das Fraunhofersche Beugungsbild eines Spaltes auf dem Negativ aufgenommen und vorschriftsmäßig vergrößert oder kopiert wird. Die Lage der schwarzen und der weißen Übergangsgrenze ist im Vergleich zum Abstand der Nebenmaxima zu vermessen, und das zugehörige Intensitätsverhältnis aus dem theoretisch bekannten Intensitätsverlauf des Beugungsbildes zu berechnen. Dieses Intensitätsverhältnis ist gleich e^{2w} ; daraus folgt w .

Für eine Aufnahme auf IG-Stat-Film im Licht der grünen Hg-Linie, vergrößert auf Agfa Lupex Hart wurde so $w = 0,3$ gemessen; bei Agfa Isopan F-Film im Rot bei $\lambda = 6200 \text{ Å}$ und Vergrößerung auf Agfa Lupex Normal war $w = 0,67$. Ferner ist bezüglich älterer photographischer Mittel aus einer Reproduktion (soweit das daraus möglich ist) in einer Arbeit von Goldberg, Luther und Weigert⁷⁾ zu entnehmen, daß die Platte Sigurd-Lichthofffrei, kopiert auf Jahr-Reproduktionsplatte, die Weichheit 0,57 hatte.

⁷⁾ Goldberg, Luther, Weigert, Z. wiss. Phot. 9, 323 (1911).

Photomechanische Platten haben eine wesentlich geringere Weichheit (0,05 etwa), sind allerdings für die meisten mit der Minimumstrahlkennzeichnung durchzuführenden Arbeiten zu empfindlich. Aber auch die Weichheit hochempfindlicher Schichten läßt sich erheblich herabsetzen, wenn das Negativ zunächst abgeschwächt und dann verstärkt wird. Als Verstärker ist der Uranverstärker brauchbar; zum Abschwächen eignet sich wohl nur der Blutlaugensalzabschwächer, der jedoch in den üblichen Konzentrationen für den vorliegenden Zweck zu plötzlich arbeitet. Empfehlenswert ist folgende Mischung: 4 cm³ Lösung (5proz.) roten Blutlaugensalzes + 100 cm³ Natriumthiosulfatlösung (1,2proz.), die weniger Blutlaugensalz als üblich enthält und langsamer wirkt. Man kontrolliere bei der Abschwächung die Minima mit der Lupe und höre auf, sobald kein Silber mehr im Minimum sichtbar ist. Das Abschwächen wird zweckmäßig unmittelbar am fixierten (gut abgespülten, aber nicht gewässerten und vor allem nicht getrockneten) Negativ vorgenommen; die sonst gelegentlich vorkommenden größeren Ungleichmäßigkeiten sind dann mit Sicherheit ausgeschlossen. Durch diesen zusätzlichen Prozeß wird naturgemäß der Schwellenwert etwas herabgesetzt, bleibt aber immer noch sehr hoch im Vergleich zu Schichten, die von vornherein bereits entsprechend hart arbeiten. Die Korngröße setzt schließlich diesem Verfahren ein Ende.

Es muß noch darauf hingewiesen werden, daß ein Herabsetzen der Weichheit allein natürlich nicht zum gewünschten Ziel führt, da flache Flanken eines Maximums beispielsweise zwar scharf aber unregelmäßig geformt werden. Wegen der mit diesen chemischen Kunstgriffen verbundenen Unsicherheit wurde bei allen mit der Minimumstrahlkennzeichnung durchgeführten Arbeiten letzten Endes auf ihre Anwendung ganz verzichtet.

Wird der Film ohne Positivprozeß unmittelbar im Komparator vermessen, so kann man auch hier eine Weichheit des Negativs w_n in analoger Weise definieren; doch hängt die Schwärzung, die als „volle Schwärzung“ zu bezeichnen wäre, unter anderem dann noch von der Beleuchtung im Komparator ab; der Unterschied zu dem Positiv liegt in der größeren Dicke der Bromsilberschicht, die auch in den stark gedeckten Teilen noch große Schwärzungsunterschiede aufweist, wenn wir nur mit hinreichender Intensität durchleuchten. Für die Beurteilung der Minima gibt es eine (übrigens individuell verschiedene) optimale Beleuchtung, die naturgemäß empirisch ermittelt werden muß. Die quantitative Erfassung dieser Dinge im Rahmen dieser Arbeit würde zu weit in Einzelheiten führen.

B) Der als Ausgangspunkt der Betrachtungen auf Seite 346 gewählte Zwang, den eine Querschnittbegrenzung auf die Richteigenschaften einer Welle ausübt, steht nur scheinbar im Widerspruch zu der bekannten Tatsache, daß beliebig scharfe Bündelung mit einem beliebig kleinen Quellgebiet erreicht werden kann, wenn dieses nur Vielpolcharakter hinreichender Vielfachheit besitzt. Durch passende Zusammenfügung hinreichend vieler Hertzscher Dipole in einem beliebig klein vorgeschriebenen Gebiet kann zwar eine Welle erregt werden, die im Fernfeld den Hauptteil ihrer Energie in einem beliebig klein vorgebbaren Winkelbereich um eine Gerade führt; aber diese Schärfe geht sofort verloren, wenn der Querschnitt des Bündels eingengt wird; der Nachweis hierzu ist in den oben angestellten Betrachtungen enthalten. Über die Quellen selbst und etwa eine Mindestgröße des Quellgebietes war oben nichts ausgesagt worden; es war lediglich mit der Benutzung der Kirchhoffschen Formel implizit vorausgesetzt worden, daß die Beobachtungsebene und die unsere Aperturblende enthaltende Hüllfläche sich in einem einfach zusammenhängenden Gebiet ohne Quellen oder andere Singularitäten befinden sollen.

Zusammenfassung

A) Eine Strahlverschärfung ist in drei Stufen möglich

1. Da bei nicht zu schwacher Intensität nur Intensitätsverhältnisse sowohl bei visueller als auch bei photographischer Beobachtung ausschlaggebend für die Strahlschärfe sind, ermöglichen hinreichend schwache Maxima geeigneter Form und Nullstellen jeder Art prinzipiell unbegrenzt hohe Strahlverschärfung.

2. Die linearen Minima vom Typ $E(\gamma) = c_1 \gamma + \dots$ führen bei gleichem Belichtungsaufwand oder gleichem Störlicht zu höherer Strahlschärfe als die quadratischen Minima oder die notwendigerweise mindestens quadratischen Maxima.

Lineare Minima treten z. B. bei dem Fresnelschen Spiegelversuch oder dem Fresnelschen Biprisma auf.

3. Das mit der gegenphasigen Belegung (27) herstellbare lineare Minimum besitzt die größte mögliche Flankensteilheit bei vorgegebener Flächenhelligkeit und ist daher allen anderen Mitteln zur Strahlverschärfung überlegen.

B) Vorteile der Minimumstrahlkennzeichnung

1. Prinzipiell unbeschränkt hohe Strahlschärfe, dadurch Steigerung der Meßgenauigkeit bis zu der durch Gerätemängel und Streulicht bedingten Grenze. Photometrie wird überflüssig.

2. Gerätemängel verursachen weniger systematische Fehler als bei anderen Strahldefinitionen.

3. Mathematisch einfachste Kennzeichnung einer Ebene oder allgemeiner Fläche im Raum mit Hilfe des Lichtes; zugehörige theoretische Rechnungen sind daher meist streng durchführbar.

4. Die Frage nach dem Lichtweg erhält durch die Minimumstrahlkennzeichnung oft erst einen eindeutigen Sinn.

5. In der Minimumbreite wird eine weitere für die theoretische Deutung wichtige Größe beobachtbar.

6. Außerhalb der gekennzeichneten Ebene ist noch in einem großen Bereich Licht vorhanden, das gleichzeitig eine Kombination mit anderen Beobachtungen (z. B. Grenzwinkel der Totalreflexion) zuläßt.

Die genannten Vorteile sind nur zum Teil in vorstehender Arbeit behandelt worden, zum anderen wurden sie bei den speziellen Anwendungen erkennbar.

Verfasser dankt Herrn Prof. Dr. Lochte-Holtgreven und Herrn Prof. Dr. Unsöld dafür, daß sie ihm die Möglichkeit zur Wiederaufnahme seiner Arbeit boten, und Herrn Prof. Dr. Lochte-Holtgreven ganz besonders für die große Bereitwilligkeit, mit der er die Wünsche des Verfassers unter schwierigsten Umständen und in einem im Wiederaufbau befindlichen Institut erfüllte.

Kiel, Institut für Experimentalphysik der Universität.

(Bei der Redaktion eingegangen am 9. April 1950.)

m
n
it

-
.
n
r
l
e

Wel
schen E
hauptur
ohne Q
wird ma
gen für
Masse,
Forderu
Entwick
wird ein
Spannun
von T_{kr}
symmet
Größe -
Faktor)
ehens so
Gl. (4,6)
lineare L
elektrost
völligen
deren ge
werden
tronenth
Der Gre
gelten vo
zwischen
geschwin
chen lau
gleich ih
dieser Th
mulierte
geschwin
V, folgen
dargest
daß näm
Geschwin
Teilchen
ter Weis
A ist die
elektrost
Ihre Lös
Potential
mit endl
Die Lad
Feldstärk
strenge L
wicklung
glied wür

Ansätze zu einer nichtlinearen Elektrodynamik

Von Karl Bechert

Inhaltsübersicht

Weil der „Elektronenradius“ r_0 aus den charakteristischen Größen e , m_0 , c der klassischen Elektronentheorie gebildet werden kann, scheint es grundsätzlich möglich, die Behauptung endlicher Ausdehnung elementarer Teilchen des Feldes mit klassischen Begriffen ohne Quantentheorie zu beschreiben. Für typische Quantenerscheinungen wie den Spin wird man natürlich ohne Quantentheorie nicht auskommen (§ 1). Es werden 10 Forderungen für eine klassische Elektrodynamik aufgestellt, die imstande sein soll, die endliche Masse, Ladung, Ausdehnung der Teilchen zu beschreiben; besonders wichtig ist die Forderung 7, daß aus der Kenntnis der Feldstärken für einen Zeitpunkt die ganze weitere Entwicklung des Feldes einschließlich der Teilchenbewegung folgen soll (§ 2). In § 3 wird eine größere Zahl algebraischer Sätze bewiesen, hauptsächlich solche über den Spannungstensor T_{kn} . Die Sätze gelten unabhängig von der physikalischen Bedeutung von T_{kn} für jeden Tensor zweiter Stufe, der in der gleichen Art wie T_{kn} aus einem antisymmetrischen Tensor zweiter Stufe F_{kl} , hier dem Feldstärkentensor, gebildet ist. Die Größe $-T_{11}V_1$ läßt sich als Ruhenergiedichte deuten; V_1 ist (bis auf einen konstanten Faktor) die Vierergeschwindigkeit (§ 4). Es gelingt damit, Impuls und Energie eines Teilchens so zu definieren, daß die richtigen Transformationseigenschaften herauskommen, Gl. (4,6). Für den Zusammenhang zwischen Viererstrom S_n und Spannungstensor wird das lineare homogene Gleichungssystem (4,9) angenommen und physikalisch begründet; für elektrostatische Modelle, auch für das Coulombsche Feld ist (4,9) identisch erfüllt. Zur völligen Bestimmung von S_n aus den Feldgrößen bleibt eine invariante Beziehung frei (4,14), deren genaue Form wohl nur die Quantentheorie liefern kann. Spezielle Ansätze für (4,14) werden in der Arbeit untersucht. Als Feldgleichungen werden die der Lorentzschen Elektronentheorie übernommen (§ 4). Die Forderung 7 ist damit erfüllt. Folgerungen aus (4,9): Der Grenzübergang $v \rightarrow c$ gibt Argumente zugunsten von (4,9). Unabhängig von (4,9) gelten vermöge der Lorentz-Transformation im Grenzfall $v \rightarrow c$ gewisse Beziehungen zwischen den Komponenten eines beliebigen Tensors (§ 5). Für ungeladene, mit Lichtgeschwindigkeit wandernde Teilchen ergibt sich als Feld die transversale Welle. Die Teilchen laufen in Richtung der Wellennormale, sie haben die Ruhmasse Null, ihr Impuls ist gleich ihrer Energie geteilt durch c (§ 5). In den nichtquantentheoretischen Gleichungen dieser Theorie stecken also alle Aussagen über Lichtquanten, die ohne das Plancksche h formuliert werden können. Nach der Theorie sind auch ungeladene Teilchen mit Unterlichtgeschwindigkeit möglich. — Die Gleichungen (4,9) zusammen mit der aus der Definition der V_n folgenden Beziehung $V_n V_n = 1$ lassen sich so auflösen, daß drei der V_n durch das letzte dargestellt werden, (6,25) oder (5,26). In § 6 wird zunächst das Trägheitsgesetz bewiesen, daß nämlich ein gradlinig gleichförmig bewegtes Teilchen ohne äußere Einwirkung seine Geschwindigkeit nicht ändert. Dann werden Lösungen der Feldgleichungen für ruhende Teilchen untersucht, wobei die in (4,14) noch unbestimmte invariante Funktion in bestimmter Weise gewählt wird: Die Invariante C des Viererstroms wird proportional A^2 gesetzt; A ist die Invariante des Spannungstensors, n muß ≥ 1 sein. $n = 1$ verlangt für rein elektrostatisches Feld die Lösung der Differentialgleichung (6,4) $\Delta u = G(u) \cdot (\text{grad } u)^2$. Ihre Lösung wird für beliebiges G und beliebig viel Dimensionen auf die Lösung der Potentialgleichung zurückgeführt. Es wird gezeigt, daß die Aufgabe nur eine Lösung mit endlicher Gesamtenergie und Gesamtladung hat, die radialsymmetrische (6,16). Die Ladung ist bei diesem Modell praktisch auf den Bereich r_0 zusammengedrängt, die Feldstärke weicht nur für $r \lesssim r_0$ merklich von der Coulombschen ab. Während die strenge Lösung endliche Werte für Masse und Ladung liefert, würde eine Reihenentwicklung „nach r_0 “ divergente Ergebnisse für diese Größen geben; jedes neue Reihenglied würde einen noch stärker divergierenden Ausdruck zu den niedrigeren Näherungen

hinzufügen. — Auch dann, wenn neben dem elektrostatischen ein magnetostatisches Feld vorhanden ist, gibt es im Fall $n = 1$ für das ruhende Teilchen nur die Lösung (6,16). Der Fall $n > 1$ könnte geeignet sein, Teilchen verschiedener Ruhmasse und entgegengesetzten Ladungsvorzeichens zu liefern. Die radialymmetrische Lösung läßt sich leicht angeben, (6,42). Bei ihr ist die Ladung noch stärker konzentriert als bei $n = 1$; Gesamtladung und Masse sind endlich. $n = \frac{3}{2}$ und $n = \frac{9}{8}$ werden besprochen; es genügt, Werte zwischen 1 und 2 zu betrachten. — Für Tensorgrößen B , die mit der Ladungsströmung wandern, gilt der Satz (7,2). Ist die Divergenz von $B V_1$ Null, so bleibt die Größe $B V_1 d\tau_3$ trotz der Ladungswanderung erhalten. — Bei fehlender äußerer Einwirkung stellt sich die gradlinig gleichförmige Bewegung ein (§ 8). Wirkt ein Feld, so bleibt die Ruhmasse entweder erhalten oder das Teilchen geht durch die Feldeinwirkung in einen anderen Teilchentyp über. Bei gleichförmiger Bewegung übt das Teilchen keine Kraft auf sich aus (§ 8). In § 9 werden in den Gleichungen (9,2), (9,4) bis (9,7) allgemeine Identitäten der Tensoranalysis angegeben; beim Einsetzen von Tensorgrößen, welche weiteren Beziehungen, z. B. Feldgleichungen genügen, werden diese Identitäten zu Erhaltungssätzen. Als Spezialfall ergibt sich der Energie-Impulserhaltungssatz der Lorentzschen Theorie. Zu jedem Erhaltungssatz kann man eine Strömungsgeschwindigkeit definieren; wandern die Grenzen eines Raumgebietes $d\tau_3$ mit dieser Geschwindigkeit, so bleibt eine gewisse, aus dem Tensor gebildete Größe bei der Bewegung zeitlich konstant (§ 9). In § 10 Bemerkungen zur wellenmechanischen Deutung dieser Theorie: Für ein einzelnes Teilchen wird versucht, die Größe $\frac{q}{e}$ als Aufenthaltswahrscheinlichkeit zu deuten, q ist die Ladungsdichte. Zerlegung in de Broglie-Wellen gibt für Impulse $< \frac{2\hbar}{r_0}$ ein Verhalten, das bei einem punktförmigen Teilchen plausibel ist: Gleichwahrscheinlichkeit für alle Impulsbeträge. Andererseits kommen Impulse, die groß sind gegen $\frac{2\hbar}{r_0}$, nicht merklich vor; das Impulsspektrum ist also nach oben praktisch „abgeschnitten“, — in relativistisch invarianter Weise, denn die ganze Theorie ist lorentzinvariant.

§ 1. Einleitung

Die vorliegende Arbeit¹⁾ gibt keine fertige Theorie, sie will auf Zusammenhänge und Fragen aufmerksam machen, deren Kenntnis für eine Lösung des Problems der Elementarteilchen vielleicht nützlich ist.

Bekanntlich kann man aus den charakteristischen Größen e , m_0 , c der klassischen Elektronentheorie eine Länge bilden, den „klassischen Elektronenradius“ (wir verwenden rationale Gaußsche Einheiten)

$$r_0 = \frac{e^2}{4\pi m_0 c^2}; \quad (1,1)$$

die Behauptung endlicher Ausdehnung elementarer Teilchen des Feldes ist also in den Begriffen der klassischen Elektrodynamik enthalten, sofern man m_0 mit einer elektromagnetischen Größe identifizieren kann.

In diesem Zusammenhang ist bemerkenswert, daß man mit den in den Feldgleichungen der Elektronentheorie vorkommenden Dimensionen, nämlich mit denen der elektrischen Feldstärke \mathcal{E} , der Ladungsdichte q , und der Lichtgeschwindigkeit c , die Dimensionen beliebiger mechanischer und auch atomphysikalischer Größen bilden, und zwar eindeutig bilden kann. So ergibt sich:

$$\frac{\mathcal{E}^2}{q^2 c^2} \sim m_0; \quad \frac{\mathcal{E}}{q} \sim r_0; \quad \frac{\mathcal{E}^2}{q^2 c} \sim \hbar; \quad \frac{\mathcal{E}^2}{q^2} \sim e^2. \quad (1,2)$$

¹⁾ In ihrem wesentlichen Inhalt ist sie im Sommer 1947 entstanden. Im Winter 1947/48 habe ich sie mit Herrn Dr. H. Marx, Wetzlar diskutieren können. Am 14. 5. 1949 habe ich auf der Tagung der Physikal. Gesellschaft von Rheinland-Pfalz die wichtigsten Ergebnisse vorgetragen.

Im Folgenden werden Forderungen zusammengestellt, die eine klassische Theorie der Elementarteilchen wohl erfüllen muß. Natürlich meine ich nicht, daß eine befriedigende Theorie elementarer Teilchen ohne Quantenbegriffe geschaffen werden könne; für die Beschreibung des Spins, für die Diskussion der Messungsmöglichkeiten, für die endgültige Formulierung der Wechselwirkungen wird man ohne Quantenbegriffe nicht auskommen. Das Problem der Masse und der Ausdehnung der Teilchen scheint aber mit Begriffen der klassischen Physik angreifbar.

§ 2. Forderungen an eine klassische Theorie elementarer Teilchen

1. Die gesamte Energie des elektromagnetischen Feldes muß für ein ruhendes Teilchen endlich sein:

$$\int T_{44} d\tau_3 \text{ endlich.} \quad (2,1)$$

Diese Forderung wird bekanntlich durch die klassische Elektrodynamik nicht erfüllt; dort ist $T_{44} \sim \mathcal{E}^2$ für ein ruhendes Teilchen (\mathcal{E} ist die elektrische Feldstärke), und es gilt $\mathcal{E} \sim \frac{1}{r^2}$; das Integral divergiert wie $\left(\frac{1}{r}\right)_{r \rightarrow 0}$. Man hat daran gedacht, durch „Abschneidevorschriften“ das Integral endlich zu machen. Das läuft darauf hinaus, nicht von $r = 0$ an, sondern von einer endlichen Länge $\approx r_0$ aus bis $r = \infty$ zu integrieren. Damit wird das „Innere“ des Teilchens aus dem Feld ausgeschlossen. Für das „Abschneiden“ gibt es keinen anderen Grund als den Wunsch, ein endliches Ergebnis zu erhalten, statt des von der Theorie gelieferten unendlichen. Es sind Theorien von kugelförmigen Teilchen mit gegebenem „Ruhradius“ proportional r_0 versucht worden. Gegen einen solchen Versuch erhebt sich die Schwierigkeit, zu erklären, warum eine solche Ladungsverteilung zusammenhält.

2. Die Gesamtladung des Teilchens muß endlich sein:

$$\int \rho d\tau_3 \text{ endlich;} \quad (2,2)$$

ρ ist die Raumladungsdichte. Betrachtet man die Teilchen als streng punktförmig, so ist die Bedingung (2,2) überflüssig, es hat dann nur Sinn, von der Gesamtladung des Teilchens zu reden.

3. In „großer“ Entfernung vom Teilchen muß das Feld des Teilchens das Coulombsche Feld sein:

$$\mathcal{E}_r \rightarrow \frac{e}{4\pi r^2} \quad (2,3)$$

für hinreichend großes r . „Groß“ wird man in diesem Zusammenhang eine Entfernung vom Teilchen nennen können, wenn sie groß ist gegenüber der charakteristischen Länge r_0 des Feldes.

4. In der Nähe des Teilchens, d. h. für $r \lesssim r_0$ wird \mathcal{E} vom Coulombfeld abweichen müssen, weil sonst die Bedingung 1. nicht erfüllt werden kann.

5. Die Ruhmasse des Teilchens muß bei der Bewegung erhalten bleiben. Da die Ruhmasse durch das Integral (2,1) über elektromagnetische Größen definiert wird, ist ihr Erhaltenbleiben in einer elektromagnetischen Theorie nicht selbstverständlich. Durch Bewegung wird das Feld verändert, es ist nicht ohne weiteres zu erwarten, daß das Integral (2,1) nach Ablauf einer Bewegung denselben Wert im Ruhssystem des Teilchens hat wie vorher.

6. Die Energie des Teilchens muß sich wie die vierte Komponente eines Vierervektors transformieren, und es muß sich ein Impuls für das Teilchen definieren lassen, der die übrigen drei Komponenten dieses Vierervektors stellt.

7. Wenn der Zustand des elektromagnetischen Feldes für einen Zeitpunkt ($t = 0$) bekannt ist, so muß die weitere Entwicklung des Feldes zwangsläufig daraus folgen, also auch die Bewegung der Teilchen. Die Lorentzsche Elektronentheorie dagegen verlangt die Kenntnis der Ladungsverteilung ρ und des Teilchenstromes $\rho \mathbf{v}$ für alle Zeiten, damit aus ihr die Entwicklung des Feldes abgeleitet werden könne. Bei Lorentz sind die Teilchen die Ursache des Feldes, ihre Bewegung muß bekannt sein, damit Voraussagen über das Feld möglich werden. Die Teilchen gehören nach Lorentz nicht selbst zum Feld, „sie sind Fremdkörper der Theorie“, hat man gesagt.

In einer konsequenten Theorie müssen aber die Teilchen mit zum Feld gerechnet werden. Durch die Vorgabe des Anfangszustandes des Feldes muß die Anfangslage und die Anfangsgeschwindigkeit der Teilchen bereits mitgegeben sein, wie es sich für einen Anfangszustand im Sinn der Mechanik gehört. Denn die rein elektromagnetische Theorie muß in ihrem Aussagenschema auch die Aussagen über die mechanischen Eigenschaften der Teilchen enthalten. Der Anfangszustand im Sinn der konsequenten Theorie, also die Vorgabe von \mathcal{E} , \mathcal{H} (\mathcal{H} ist die magnetische Feldstärke), muß den Anfangszustand für die Teilchen auch im Sinn der Mechanik festlegen. Da das vorhandene Feld die auf die Teilchen wirkenden Kräfte mitenthält, so muß die weitere Bewegung der Teilchen und die Feldveränderung durch den Anfangszustand völlig bestimmt sein.

Die Forderung 7 sieht besonders unangenehm aus. In der gewöhnlichen Elektrodynamik ist sie erfüllt: Gibt man \mathcal{E} , \mathcal{H} im ganzen Raum vor, so sind die partiellen zeitlichen Differentialquotienten $\dot{\mathcal{E}}$, $\dot{\mathcal{H}}$ aus den Feldgleichungen bekannt, wenn die übrigen Feldgrößen \mathcal{B} , \mathcal{J} , \mathcal{D} als Funktionen von \mathcal{H} , \mathcal{E} bekannt sind; durch Differentiation der Feldgleichungen nach t erhält man dann auch $\dot{\mathcal{E}}$, $\dot{\mathcal{H}}$ und alle höheren zeitlichen Differentialquotienten; \mathcal{E} , \mathcal{H} sind damit für alle folgenden Zeitpunkte im ganzen Raum aus der Anfangsbedingung bestimmt.

8. Für die Bewegung geladener Teilchen muß mindestens in erster Näherung die Lorentzsche Bewegungsgleichung gelten mit der Lorentzischen Kraft:

$$\mathbf{K} = q \cdot \mathcal{E} + q \left[\frac{\mathbf{v}}{c} \mathcal{H} \right] \quad (2,4)$$

auf die Volumeneinheit der Ladung.

9. Die Rückwirkung des Feldes auf ein geladenes bewegtes Teilchen muß wenigstens in erster Näherung durch die Lorentzsche Strahlungsdämpfung gegeben sein.

10. Die Theorie muß lorentzinvariant sein.

Für instabile Teilchen bleiben die Forderungen 1., 2., 6., 7., 10., für ungeladene stabile Teilchen kommt noch 5. dazu.

§ 3. Allgemeine algebraische Sätze, welche für jede elektrodynamische Theorie gelten

Wir benutzen die vierdimensionale Schreibweise, in der man die Kovarianzeigenschaften gegenüber Lorentztransformationen unmittelbar sehen kann. Die Feldgrößen sind zusammengefaßt in dem antisymmetrischen Tensor F_{kl} ; Komponenten nach x, y, z, ict bezeichnen wir mit den Indices 1, 2, 3, 4, über gleiche Indizes in Produkten und über gleiche Indizes am selben Tensor ist in

allen Formeln zu summieren.

$$F_{kl} = -F_{lk}; \quad F_{kl} \equiv \begin{pmatrix} 0 & \mathfrak{H}_3 & -\mathfrak{H}_2 & -i\mathfrak{E}_1 \\ -\mathfrak{H}_3 & 0 & \mathfrak{H}_1 & -i\mathfrak{E}_2 \\ \mathfrak{H}_2 & -\mathfrak{H}_1 & 0 & -i\mathfrak{E}_3 \\ i\mathfrak{E}_1 & i\mathfrak{E}_2 & i\mathfrak{E}_3 & 0 \end{pmatrix}; \quad (3,1)$$

die Koordinaten sind:

$$x_i \equiv \{x, y, z, ict\}. \quad (3,2)$$

Der Vierersturm heißt S_k :

$$S_k \equiv \frac{q}{c} \{v_x, v_y, v_z, ic\}. \quad (3,3)$$

Das Differential $d\tau$ der Eigenzeit definieren wir wie üblich:

$$(ic d\tau)^2 \equiv dx_k dx_k, \quad (3,4)$$

und führen einen der Vierergeschwindigkeit proportionalen Vierervektor V_k ein:

$$V_k \equiv \frac{dx_k}{ic d\tau} = \frac{1}{\sqrt{1-\beta^2}} \left\{ -\frac{iv_x}{c}, -\frac{iv_y}{c}, -\frac{iv_z}{c}, 1 \right\}; \quad \beta \equiv \frac{|v|}{c}. \quad (3,5)$$

Wir werden V_k im folgenden die Vierergeschwindigkeit nennen.

Für reelle Bewegung sind die drei ersten Komponenten von V_k rein imaginär, während die vierte reell ist.

Es gilt:

$$V_k V_k = 1. \quad (3,6)$$

Solange man V_k durch (3,5) einführt, ist (3,6) eine Identität. Berechnet man aber V_k aus Feldgrößen, und das werden wir tun, so bedeutet dies, daß diese aus den Feldgrößen gewonnene Größe V_k mit (3,5) identifiziert wird. Dann ist (3,6) eine Bedingung, ebenso wie die genannten Realitätsforderungen an die V_k dann Bedingungen darstellen.

Unser erster Satz ist eine Trivialität. Er lautet:

1) $\frac{S_k}{V_k}$ ist eine Invariante.

Das bedeutet zugleich, daß $\frac{S_k}{V_k}$ vom Index k unabhängig ist. Also:

$$S_k = \frac{S_4}{V_4} V_k; \quad (3,7)$$

für $k = 1, 2, 3$ folgt dies aus den Definitionen (3,3), (3,5) für $k = 4$ ist es eine Identität. Es gilt:

$$\frac{S_4}{V_4} = \frac{S_k}{V_k} = iq \sqrt{1-\beta^2} \equiv C = \text{Invariante}. \quad (3,8)$$

Die Aussage „ $q \sqrt{1-\beta^2}$ eine Invariante“ ist nicht neu.

Wir betrachten den symmetrischen Tensor T_{kn} ($\delta_{kn} = 1$, wenn $k = n$; $\delta_{kn} = 0$ sonst):

$$T_{kn} = F_{ki} F_{ni} - \frac{1}{4} \delta_{kn} F_{js} F_{js}; \quad T_{kn} = T_{nk}. \quad (3,9)$$

In der klassischen Elektrodynamik werden seine Komponenten gedeutet als Maxwell'sche Spannungen, als Impuls- und Energiedichte des Feldes. Wir führen die Bezeichnungen ein:

$$\begin{aligned} \text{I} &\equiv \mathfrak{E}_1^2 + \mathfrak{H}_1^2; \quad \text{II} \equiv \mathfrak{E}_2^2 + \mathfrak{H}_2^2; \quad \text{III} \equiv \mathfrak{E}_3^2 + \mathfrak{H}_3^2; \\ W &\equiv \frac{1}{2} (\mathfrak{E}^2 + \mathfrak{H}^2); \quad \mathfrak{S} \equiv c [\mathfrak{E} \mathfrak{H}]. \end{aligned} \quad (3,10)$$

Dann gilt:

$$\left. \begin{aligned} T_{11} &= W - I; \quad T_{22} = W - II; \quad T_{33} = W - III; \quad T_{44} = -W; \\ T_{12} &= -(\mathfrak{E}_1 \mathfrak{E}_2 + \mathfrak{H}_1 \mathfrak{H}_2); \quad T_{13} = -(\mathfrak{E}_1 \mathfrak{E}_3 + \mathfrak{H}_1 \mathfrak{H}_3); \\ T_{23} &= -(\mathfrak{E}_2 \mathfrak{E}_3 + \mathfrak{H}_2 \mathfrak{H}_3); \quad T_{14} = \frac{i \mathfrak{E}_1}{c}; \quad T_{24} = \frac{i \mathfrak{E}_2}{c}; \quad T_{34} = \frac{i \mathfrak{E}_3}{c}. \end{aligned} \right\} (3,11)$$

Wir merken an, daß:

$$F_{js} F_{js} = 2 (\mathfrak{H}^2 - \mathfrak{E}^2). \quad (3,12)$$

Der Tensor T_{kn} hat eine Reihe merkwürdiger Eigenschaften. Als erste nennen wir die bekannte Eigenschaft:

$$II) \quad T_{44} = 0,$$

die aus (3,9) oder (3,11) leicht bestätigt werden kann.

Es gelten die Beziehungen:

$$\begin{aligned} III) \quad T_{11} T_{22} - T_{33} T_{44} &= I \cdot II = T_{12}^2 - T_{34}^2; \quad T_{22} T_{33} - T_{11} T_{44} = II \cdot III \\ &= T_{23}^2 - T_{14}^2; \quad T_{33} T_{11} - T_{22} T_{44} = III \cdot I = T_{31}^2 - T_{24}^2. \end{aligned}$$

In der Tat ist:

$$T_{11} T_{22} - T_{33} T_{44} = (W - I)(W - II) + (W - III)W = 2W^2 - W(I + II + III) + I \cdot II.$$

Nach (3,10) ist aber:

$$I + II + III = 2W, \quad (3,13)$$

so daß in der vorigen Gleichung rechts nur $I \cdot II$ übrigbleibt. Andererseits ist:

$$I \cdot II = (\mathfrak{E}_1^2 + \mathfrak{H}_1^2)(\mathfrak{E}_2^2 + \mathfrak{H}_2^2) = (\mathfrak{E}_1 \mathfrak{E}_2 + \mathfrak{H}_1 \mathfrak{H}_2)^2 + (\mathfrak{E}_1 \mathfrak{H}_2 - \mathfrak{E}_2 \mathfrak{H}_1)^2 = T_{12}^2 - T_{34}^2.$$

Damit ist die erste Gleichung III) bewiesen, die übrigen beweist man ebenso.

Außerdem gilt:

$$IVa) \quad T_{12} T_{42} + T_{13} T_{43} = T_{11} \cdot I; \quad T_{23} T_{43} + T_{21} T_{41} = T_{24} \cdot II;$$

$$T_{31} T_{41} + T_{32} T_{42} = T_{34} \cdot III; \quad \text{und:}$$

$$IVb) \quad T_{21} T_{31} + T_{24} T_{34} = -T_{23} \cdot I; \quad T_{32} T_{12} + T_{34} T_{14} = -T_{31} \cdot II;$$

$$T_{13} T_{23} + T_{14} T_{24} = -T_{12} \cdot III.$$

Diese Gleichungen bestätigt man leicht durch direkte Ausrechnung. Aus III) und IV) folgt der Satz:

V) Alle einander zugeordneten zweireihigen Unterdeterminanten der Determinante $\|T_{kn}\|$ sind einander gleich. Die Unterdeterminanten sind einschließlich ihrer Vorzeichen gemeint. Unter der zugeordneten Unterdeterminante verstehen wir die Restdeterminante, die übrigbleibt, wenn man die Zeilen und Spalten der betrachteten Unterdeterminante aus der Determinante $\|T_{kn}\|$ streicht. Die Vorzeichen bestimmen sich so: Besteht die betrachtete Unterdeterminante aus den Zeilen i, k und den Spalten l, m , so ist sie gleich der Restdeterminante, multipliziert mit $(-)^{i+k+l+m}$.

Die Gleichungen III) enthalten bereits 3 Aussagen über solche Unterdeterminanten. Die übrigen kann man nach folgendem Schema bestätigen:

$$T_{11} T_{23} - T_{12} T_{13} - (-)^{1+3+1+2} (T_{23} T_{44} - T_{24} T_{34}) = (W - I + T_{44}) T_{23} - (T_{12} T_{13} + T_{24} T_{34})$$

$$= -I \cdot T_{23} + T_{23} \cdot I = 0,$$

nach (3,11) und der ersten Gleichung IVb). Der andere bei der Bestätigung von V) vorkommende Typ ist:

$$T_{11} T_{24} - T_{12} T_{14} - (-)^{1+4+1+2} (T_{23} T_{34} - T_{24} T_{33}) = (W - I) T_{24} + T_{23} T_{34} - T_{24} \cdot II$$

$$- T_{23} T_{34} + T_{24} (W - III) = T_{24} (2W - I - II - III) = 0,$$

wegen (3,11), der 2. Gleichung IVa) und (3,13).

Wir untersuchen die Invarianten des Tensors T_{kn} . Die erste der Invarianten, T_{ii} , ist nach II) gleich Null. Für die nächste gilt der merkwürdige Satz:

$$\text{VI)} \quad \begin{vmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{vmatrix} + \begin{vmatrix} T_{11} & T_{13} \\ T_{13} & T_{33} \end{vmatrix} + \begin{vmatrix} T_{11} & T_{14} \\ T_{14} & T_{44} \end{vmatrix} + \begin{vmatrix} T_{22} & T_{23} \\ T_{23} & T_{33} \end{vmatrix} + \begin{vmatrix} T_{22} & T_{24} \\ T_{24} & T_{44} \end{vmatrix} + \begin{vmatrix} T_{33} & T_{34} \\ T_{34} & T_{44} \end{vmatrix} \\ = -2 T_{i1} T_{i1} = -2 T_{i2} T_{i2} = -2 T_{i3} T_{i3} = -2 T_{i4} T_{i4}.$$

Der Satz bedeutet also auch, daß die Quadratsummen der Glieder aller Zeilen oder Spalten der Determinante $\|T_{kn}\|$ einander gleich sind. Es folgt:

$$\text{VIa)} \quad \frac{1}{4} T_{ni} T_{ni} = T_{n1} T_{n1} = T_{n2} T_{n2} = T_{n3} T_{n3} = T_{n4} T_{n4}.$$

Wir wollen diese Invariante mit A^2 bezeichnen. Sie hat den Wert:

$$\text{VIb)} \quad \frac{1}{4} T_{ni} T_{ni} = A^2 = T_{n4} T_{n4} = W^2 - \frac{\mathfrak{E}^2}{c^2}.$$

Wir setzen fest:

$$A = -\sqrt{T_{n4} T_{n4}}. \quad (3,14)$$

VIc) A^2 ist immer ≥ 0 .

Denn man kann A^2 so umformen:

$$A^2 = \frac{1}{4} \{(\mathfrak{E}^2 + \mathfrak{H}^2)^2 - 4 [\mathfrak{E} \mathfrak{H}]^2\} = \frac{1}{4} \{(\mathfrak{H}^2 - \mathfrak{E}^2)^2 + 4 (\mathfrak{E} \mathfrak{H})^2\}. \quad (3,15)$$

Es ist also immer $W \geq \frac{|\mathfrak{E}|}{c}$. Weiter folgt:

VIa) $T_{ni} T_{ni}$ und damit A kann nur dann Null sein, wenn $|\mathfrak{E}| = |\mathfrak{H}|$ und $\mathfrak{E} \perp \mathfrak{H}$; in diesem Fall ist auch $W = \frac{|\mathfrak{E}|}{c}$. In allen anderen Fällen ist $W > \frac{|\mathfrak{E}|}{c}$.

Zum Beweis von VI) faßt man die Unterdeterminanten gemäß III) paarweise zusammen. Es entsteht:

$$2(T_{11} T_{22} - T_{12}^2) + 2(T_{11} T_{33} - T_{13}^2) + 2(T_{11} T_{44} - T_{14}^2),$$

was wegen II) sich in die Form $-2 T_{i1} T_{i1}$ setzen läßt. Man hätte aber auch so ordnen können:

$$2(T_{22} T_{11} - T_{12}^2) + 2(T_{22} T_{33} - T_{23}^2) + 2(T_{22} T_{44} - T_{24}^2),$$

was wegen II) die Form $-2 T_{i2} T_{i2}$ annimmt. Beim Ordnen nach T_{33} ergibt sich $-2 T_{i3} T_{i3}$, beim Ordnen nach T_{44} schließlich $-2 T_{i4} T_{i4}$.

Für die Invariante dritter Ordnung gilt:

$$\text{VII)} \quad \begin{vmatrix} T_{11} & T_{12} & T_{13} \\ T_{12} & T_{22} & T_{23} \\ T_{13} & T_{23} & T_{33} \end{vmatrix} + \begin{vmatrix} T_{11} & T_{12} & T_{14} \\ T_{12} & T_{22} & T_{24} \\ T_{14} & T_{24} & T_{44} \end{vmatrix} + \begin{vmatrix} T_{11} & T_{13} & T_{14} \\ T_{13} & T_{33} & T_{34} \\ T_{14} & T_{34} & T_{44} \end{vmatrix} + \begin{vmatrix} T_{22} & T_{23} & T_{24} \\ T_{23} & T_{33} & T_{34} \\ T_{24} & T_{34} & T_{44} \end{vmatrix} = 0.$$

Man beweist das durch Ausnutzen von V) und II), indem man die erste Determinante nach den Gliedern der ersten Zeile entwickelt, die zweite und dritte nach den Gliedern der zweiten Zeile, die vierte nach den Gliedern der dritten Zeile.

Die letzte Invariante von T_{kn} , die Determinante $\|T_{kn}\|$ hat den Wert:

$$\text{VIII)} \quad \begin{vmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{12} & T_{22} & T_{23} & T_{24} \\ T_{13} & T_{23} & T_{33} & T_{34} \\ T_{14} & T_{24} & T_{34} & T_{44} \end{vmatrix} = \left(\frac{1}{4} T_{ni} T_{ni}\right)^2 = \frac{1}{16} T_{ni} T_{ni} T_{lj} T_{lj} = A^4 = \left(W^2 - \frac{\mathfrak{E}^2}{c^2}\right)^2.$$

Zum Beweis von VIII) braucht man den Satz:

$$\text{IX)} \quad T_{kl} T_{kj} = A^2 \cdot \delta_{lj}.$$

Die einander entsprechenden Glieder zweier Parallelreihen von $\|T_{kn}\|$ geben demnach miteinander multipliziert und addiert 0; für $l = j$ erhält man den Satz VIa) als Spezialfall von IX).

Der Beweis von IX) gelingt durch Ausnützen von II) und V). Zum Beispiel ist:

$$T_{11}T_{12} + T_{12}T_{22} + T_{13}T_{23} + T_{14}T_{24} = T_{12}(T_{11} + T_{22}) + T_{13}T_{23} + T_{14}T_{24} \\ = -T_{12}(T_{33} + T_{44}) + T_{13}T_{23} + T_{14}T_{24} = 0,$$

weil nach V):

$$\begin{vmatrix} T_{12} & T_{13} \\ T_{23} & T_{33} \end{vmatrix} = (-)^{2+3+1+3} \begin{vmatrix} T_{12} & T_{24} \\ T_{14} & T_{44} \end{vmatrix}.$$

Zum Beweis von VIII) entwickle man die Determinante nach zweireihigen Unterdeterminanten, etwa der ersten 2 Zeilen. Wegen V) entsteht eine Summe von Quadraten:

$$\begin{vmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{vmatrix}^2 + \begin{vmatrix} T_{11} & T_{13} \\ T_{12} & T_{23} \end{vmatrix}^2 + \begin{vmatrix} T_{11} & T_{14} \\ T_{12} & T_{24} \end{vmatrix}^2 + \begin{vmatrix} T_{12} & T_{13} \\ T_{22} & T_{23} \end{vmatrix}^2 + \begin{vmatrix} T_{12} & T_{14} \\ T_{22} & T_{24} \end{vmatrix}^2 + \begin{vmatrix} T_{13} & T_{14} \\ T_{23} & T_{24} \end{vmatrix}^2 \\ = T_{11}^2(T_{22}^2 + T_{23}^2 + T_{24}^2) + T_{12}^2(T_{12}^2 + T_{13}^2 + T_{14}^2 + T_{22}^2 + T_{23}^2 + T_{24}^2) \\ + T_{13}^2(T_{22}^2 + T_{23}^2) - 2\{T_{11}T_{22}T_{12}^2 + T_{11}T_{23}T_{13}T_{12} + T_{11}T_{24}T_{14}T_{12} + T_{13}T_{23}T_{12}T_{13} \\ + T_{12}T_{24}T_{22}T_{14} + T_{13}T_{24}T_{23}T_{14}\};$$

der Ausdruck: $2\{\}$ läßt sich umformen:

$$2 \cdot \{\} = T_{11}T_{12}(T_{12}T_{22} + T_{13}T_{23} + T_{14}T_{24}) + T_{21}T_{22}(T_{11}T_{21} + T_{13}T_{23} + T_{14}T_{24}) \\ + T_{31}T_{32}(T_{11}T_{21} + T_{12}T_{22} + T_{14}T_{24}) + T_{41}T_{42}(T_{11}T_{21} + T_{12}T_{22} + T_{13}T_{23}) \\ = -T_{11}^2T_{12}^2 - T_{12}^2T_{22}^2 - T_{13}^2T_{23}^2 - T_{14}^2T_{24}^2,$$

wegen IX). So wird aus der Determinante $\|T_{kn}\|$ wegen IX) schließlich:

$$(T_{11}^2 + T_{12}^2 + T_{13}^2 + T_{14}^2)(T_{21}^2 + T_{22}^2 + T_{23}^2 + T_{24}^2) = A^2 \cdot A^2 = A^4.$$

IX) könnte zu der Meinung verleiten, daß man die Größen $\frac{T_{kn}}{A}$ als Richtungscosinus im Vierdimensionalen verwenden könnte; das ist aber unmöglich, weil die T_{kn} nicht linear unabhängig sind, wie II) zeigt.

Wir bilden jetzt formal eine Säkulargleichung für den Tensor T_{kn} , schreiben also die Gleichung hin, welche die Hauptachsen des Tensors T_{kn} gibt, wenn man ihn als Fläche zweiten Grades im vierdimensionalen Raum auffaßt. Im nächsten Paragraphen werden wir die Säkulargleichung physikalisch deuten; hier wollen wir nur formale Beziehungen aufstellen. Die Säkulargleichung für T_{kn} lautet:

$$\|T_{kn} - B \cdot \delta_{kn}\| = 0; \quad (3,16)$$

wegen VIII), VII), VIa), VI), II) wird daraus: $A^4 - 2B^2 \cdot A^2 + B^4 = (A^2 - B^2)^2 = 0$. Daraus folgt:

X) Die Säkulargleichung (3,16) hat die Lösung:

$$B^2 = A^2 = \frac{1}{4} T_{ni} T_{ni}; \quad B = \pm A. \quad (3,17)$$

B ist (bei reellem \mathfrak{E} , \mathfrak{S}) immer reell und nur im Falle: $|\mathfrak{E}| = |\mathfrak{S}|$, $\mathfrak{E} \perp \mathfrak{S}$ gleich Null.

XI) Die Matrix $\|T_{kn} - B \delta_{kn}\|$ hat den Rang 2; alle Unterdeterminanten mit 3 Zeilen und Spalten sind identisch null, wenn B der Bleichung (3,17) genügt, also Lösung von (3,16) ist.

Der Beweis beruht auf den Eigenschaften IX) und V).

XII) Die dreireihigen Unterdeterminanten Δ_{ii} der Determinante $\|T_{ik}\|$ sind gleich den zugeordneten Elementen T_{ii} der Determinante $\|T_{ik}\|$, multipliziert mit A^2 . Die Unterdeterminanten sind dabei wieder mit Vorzeichen gemeint, das Vorzeichen ist $(-)^{i+j}$. In einer Formel:

$$\Delta_{ii} = T_{ii} \cdot A^2. \quad (3,18)$$

Zum Beispiel ist:

$$\begin{vmatrix} T_{12} & T_{23} & T_{34} \\ T_{13} & T_{23} & T_{34} \\ T_{14} & T_{34} & T_{44} \end{vmatrix} = -T_{12} \cdot A^2.$$

Beweis am einfachsten durch Entwicklung der Determinante $\Delta \equiv \|T_{ik}\|$. Es ist: $T_{ij} \Delta_{ii} = \delta_{ji} \cdot \Delta = \delta_{ji} \cdot A^4$; andererseits: $T_{ij} \Delta_{ii} = \delta_{ji} A^2$ nach IX); daher: $T_{ij}(\Delta_{ii} - T_{ii} A^2) = 0$.

Hält man l fest, so sind dies wegen $j = 1, 2, 3, 4$ vier lineare homogene Gleichungen für die 4 Unbekannten $A_{il} - T_{il} A^2$. Wenn $A \neq 0$, d. h. $A \neq 0$, so folgt die Gleichung (3,18) für festes l ; der Schluß läßt sich für jedes l machen, womit (3,18) bewiesen ist für $A \neq 0$. Für $A = 0$ sind nach dem Satz XI) wegen $B = \pm A$ alle 3reihigen Unterdeterminanten von Δ gleich Null, was die Gültigkeit von (3,18) auch für $A = 0$ sicherstellt.

XIII) A^2 ist die einzige von Null verschiedene Invariante des Tensors T_{ik} . Der Beweis folgt aus den Sätzen II), VI), VIIa), VIb), VII), VIII).

XIV) Die Invarianten des Tensors F_{ik} sind $\mathfrak{F}^2 - \mathfrak{E}^2$ und $(\mathfrak{E} \mathfrak{F})^2$, wie man leicht durch direkte Rechnung bestätigt. Natürlich setzt sich A^2 , die Invariante von T_{ik} , aus diesen beiden Invarianten zusammen, vergleiche (3,15).

In den vorigen Sätzen haben wir die Invarianten $\frac{S_k}{V_k} = \frac{S_4}{V_4} = C$ und $A^2 = \frac{1}{4} T_{ni} T_{ni} = T_{n4} T_{n4}$ kennengelernt. Daß diese Größen wirklich Invarianten gegenüber Lorentztransformationen sind, kann man leicht direkt nachrechnen. Die Lorentztransformation schreiben wir allgemein:

$$x_k = x'_i \alpha_{ki} \quad (3,19)$$

und erhalten:

$$\frac{S_k}{V_k} = \frac{S_4}{V_4} = C = \frac{S'_i \alpha_{ki}}{V'_n \alpha_{kn}}; \quad (3,20)$$

da aber identisch gilt:

$$\frac{S'_k}{V'_k} = \frac{S'_4}{V'_4} = C',$$

so folgt weiter:

$$C = \frac{C' V'_i \alpha_{ki}}{V'_n \alpha_{kn}} = C', \quad (3,21)$$

was, wie erwähnt, eine Trivialität ist. Die Invarianz von $T_{n4} T_{n4}$ ist nicht so selbstverständlich; sie beruht auf dem Satz IX). Es ist:

$$A^2 = T_{n4} T_{n4} = T_{ij} \alpha_{ni} \alpha_{4j} T'_{km} \alpha_{nk} \alpha_{4m} = \delta_{kl} T'_{ij} T'_{km} \alpha_{4j} \alpha_{4m} = T'_{kj} T'_{km} \alpha_{4j} \alpha_{4m} = A'^2. \quad (3,22)$$

Zum Schluß machen wir darauf aufmerksam, daß wir in diesem Paragraphen keinerlei Elektrodynamik getrieben haben, wir haben keine Feldgleichungen aufgestellt, vielmehr nur Gleichungen und Aussagen abgeleitet, die aus der Definition der betrachteten Größen folgen. Da diese Größen wohl in jeder Elektrodynamik eine Rolle spielen, gelten die Beziehungen für jede Elektrodynamik, unabhängig von den Zusammenhängen, welche durch die Feldgleichungen gefordert werden. Die für T_{kn} bewiesenen Sätze gelten rein algebraisch für jeden symmetrischen Tensor T_{kn} , der nach Art von (3,9) aus einem antisymmetrischen Tensor F_{kj} gebildet ist. Von den Realitätseigenschaften der F_{kj} (F_{k4} rein imaginär für $k \neq 4$, alle andern reell) sind nur die Sätze VIc, d) abhängig, die übrigen gelten für beliebige komplexe F_{kj} und T_{kn} . Sie scheinen weder in der Physik noch in der Mathematik bisher bekannt zu sein.

§ 4. Ansatz zu einem Begriffs- und Gleichungsschema für eine konsequente Theorie geladener Teilchen (ohne Beachtung von Quantenerscheinungen)

Die Forderungen 1 und 2 von § 2 bedeuten: Damit die Gesamtladung des Teilchens endlich bleibt, muß ϱ für kleine Abstände r (vom Teilchenmittelpunkt $r = 0$) sich wie $r^{-\delta}$ verhalten, mit $\delta < 3$. Den Zusammenhang $\text{div } \mathfrak{D} \sim \varrho$ wird man beibehalten. Das bedeutet $\mathfrak{D}_r \sim r^{-\varepsilon}$ und $\varrho \sim \frac{1}{r^3} \cdot \frac{\partial}{\partial r} (r^{3-\varepsilon}) = (2-\varepsilon) r^{-\varepsilon-1}$; also $\varepsilon < 2$. Das Coulombsche Feld ist ein ausgearteter Fall; dort ist $\varepsilon = 2$ und ϱ überall null außer am Ort der punktförmigen Ladung. Dort befindet sich end-

liche Ladung auf beliebig kleinem Raum, denn das Integral $\int \mathfrak{D}_n d\sigma \sim r^{-\epsilon+2}$ ist endlich. Man sieht: Für $\epsilon < 2$ kommt das vernünftiger Ergebnis heraus, daß in beliebig kleinem Raum um $r=0$ auch nur beliebig wenig Ladung sein kann. Unsere Theorie muß liefern, daß die Ladung des Teilchens im wesentlichen auf ein Gebiet vom Radius r_0 um $r=0$ beschränkt ist.

Wir versuchen, den Zusammenhang $W \sim \mathfrak{E}^2$ für ein ruhendes freies Teilchen beizubehalten, den die klassische Elektrodynamik behauptet. Damit die gesamte elektromagnetische Energie des Teilchens endlich bleibt, muß \mathfrak{E} sich für kleine r verhalten wie $r^{-\eta}$, mit $\eta < 3/2$. Dann ist in beliebig kleinem Raum um $r=0$ auch nur

beliebig wenig Energie enthalten; denn es wird: $\int_0^r \mathfrak{E}^2 d\tau_3 \sim \int_0^r r^{-2\eta+2} dr = \frac{r^{3-2\eta}}{3-2\eta}$.

Aber auch die elektromagnetische Energie muß im wesentlichen auf einen Bereich der Ausdehnung r_0 um $r=0$ konzentriert sein.

In der Bornschen Elektrodynamik²⁾ ist $\mathfrak{D} \sim r^{-2}$ wie in der klassischen Elektrodynamik, so daß die Ladungsdichte in derselben Weise für $r=0$ singulär wird, wie beim Coulombschen Feld. Dagegen gibt Born den Zusammenhang $\mathfrak{D} = \mathfrak{E}$ auf, den die klassische Elektronentheorie forderte. Er kann so erreichen, daß die gesamte elektromagnetische Energie endlich bleibt. Wir werden überall $\mathfrak{D} = \mathfrak{E}$, $\mathfrak{B} = \mathfrak{H}$ setzen, wie es die Lorentzsche Theorie tat.

In unserer Auffassung ist ρ im allgemeinen überall von null verschieden; es ist deshalb auch jedem Raumpunkt ein bestimmter Wert der Vierergeschwindigkeit V_4 zugeordnet. Dieser Umstand ist von Bedeutung für einen Schluß, den wir alsbald ziehen wollen. Da nach unserer Lösung für geladene Teilchen die Raumladung praktisch auf die nächste Umgebung von $r=0$ beschränkt ist (siehe § 6), so haben wir im Groben immer noch das Ergebnis praktisch punktförmiger Teilchen.

Die Forderung 6 von § 2 verlangt zunächst, zu untersuchen, wie sich $\int T_{44} d\tau_3$ bei einer Lorentztransformation verhält. Das Ruhssystem des Teilchens wollen wir K nennen, ein gegen K momentan mit der Geschwindigkeit v bewegtes System K' . Die momentane Bewegungsrichtung von K' gegen K sei die gemeinsame x_1 -Richtung beider Systeme. Wir nehmen an, und dies ist eine Einschränkung über den Charakter von Teilchen, die wir hier zunächst betrachten wollen, daß das Feld des Teilchens im Ruhssystem rein elektrisch sei. Diese Annahme wird übrigens, meist stillschweigend, in allen bisherigen klassischen Theorien geladener Teilchen gemacht. Der Spin des Teilchens mit seinen magnetischen Wirkungen ist eine Quantenerscheinung, wie die Größe $1/2 \hbar$ des Spin-Impulses zeigt; er hängt wohl mit der Unschärfe zusammen und muß sich bei einer Übertragung ins Quantentheoretische ergeben.

Das Schema der Lorentztransformation lautet:

	x'_1	x'_2	x'_3	x'_4
x_1	$\frac{1}{\sqrt{1-\beta^2}}$	0	0	$\frac{-i\beta}{\sqrt{1-\beta^2}}$
x_2	0	1	0	0
x_3	0	0	1	0
x_4	$\frac{i\beta}{\sqrt{1-\beta^2}}$	0	0	$\frac{1}{\sqrt{1-\beta^2}}$

(4,1)

²⁾ M. Born, Proc. Roy. Soc. London, Ser. A., 143, 410 (1934).

Eine bekannte Überlegung zeigt, daß $\int T_{44} d\tau_3$ sich nicht wie die Energie transformiert. Wir bilden den Ausdruck für das gegen das Teilchen bewegte System K' und beachten, daß die Integration bei festem Zeitpunkt, also hier bei festem x'_4 , gemeint ist. Die Koeffizienten des Schemas (4,1) nennen wir α_{ik} und erhalten

$$\begin{aligned} (\int T'_{44} d\tau'_3)_{x'_4 = \text{const}} &= (\int T_{1j} \alpha_{14} \alpha_{j4} d\tau'_3)_{x'_4 = \text{const}} \\ &= (\int (T_{11} \alpha_{14}^2 + T_{44} \alpha_{44}^2) d\tau'_3)_{x'_4 = \text{const}}. \end{aligned} \quad (4,2)$$

(4,1) lehrt, daß für $x' = \text{const}$ gilt: $dx_1 = \frac{dx'_1}{\sqrt{1-\beta^2}}$; also: $d\tau'_3 = d\tau_3 \cdot \sqrt{1-\beta^2}$
 $= \frac{d\tau_3}{\alpha_{44}}$; $x'_4 = \text{const}$ bedeutet nicht, daß auch x_4 const, es ist vielmehr:

$$x'_4 = \frac{-i\beta x_1 + x_4}{\sqrt{1-\beta^2}}; \quad x_4 = i\beta x_1 + x'_4 \cdot \sqrt{1-\beta^2}. \quad (4,3)$$

Das Feld in K ist aber statisch, wir interessieren uns nur für die Transformation des Eigenfeldes des Teilchens, das wir als rein elektrisch, also statisch voraussetzen wollten; in K ist \mathbf{v} überall Null, die Feldgrößen hängen demnach nicht von x_4 ab. Wären sie von x_4 abhängig, so müßte man bei der Integration über τ_3 den Beitrag jedes Raumpunktes innerhalb τ_3 zu einem Zeitpunkt x_4 nehmen, der nach (4,3) zum fest vorgegebenen x'_4 und zum betrachteten Raumpunkt (x_1) gehört. Weil aber die Feldgrößen in K von x_4 unabhängig sind, können wir in K über alle Raumpunkte innerhalb τ_3 bei festem x_4 integrieren und erhalten:

$$(\int T'_{44} d\tau'_3)_{x'_4 = \text{const}} = \left(\int \frac{-\beta^2 T_{11} + T_{44}}{\sqrt{1-\beta^2}} d\tau_3 \right)_{x_4 = \text{const}} \quad (4,4)$$

Würde $\int T_{44} d\tau_3$ die Rolle der Teilchenenergie spielen, so müßte das Glied mit T_{11} verschwinden, was es wegen der Bedeutung von T_{11} offensichtlich nicht tut, siehe (3,11) und (3,10).

Die Auffassung kontinuierlicher Verteilung von ρ und V_k gibt uns die Möglichkeit, diese Schwierigkeit in einfacher Weise zu lösen. Im Ruhesystem K ist überall $\mathbf{v} = 0$, also $V_4 = 1$; das bedeutet: $V_n = \delta_{n4}$. Außerdem ist $T_{n4} = T_{44} \cdot \delta_{n4}$, weil das Feld in K rein elektrisch, also $\mathfrak{H} = 0$ ist. Mit den V_n kann man leicht einen Ausdruck bilden, der im Ruhesystem in $\int T_{44} d\tau_3$ übergeht und trotzdem die richtigen Transformationseigenschaften hat.

Wir zeigen zunächst, daß $(\int T'_{4l} V'_l d\tau'_3)_{x'_4 = \text{const}}$ eine Invariante ist. Es gilt:

$$\begin{aligned} (\int T'_{4l} V'_l d\tau'_3)_{x'_4 = \text{const}} &= (T_{nj} \alpha_{n4} \alpha_{jl} V_4 \alpha_{4l} d\tau'_3)_{x'_4 = \text{const}} \\ &= (\int T_{nj} \alpha_{n4} \delta_{j4} V_4 d\tau'_3)_{x'_4 = \text{const}} = (\int T_{n4} \alpha_{n4} V_4 \delta_{44} d\tau'_3)_{x'_4 = \text{const}} \\ &= (\int T_{44} \delta_{n4} \alpha_{n4} V_4 d\tau'_3)_{x'_4 = \text{const}} = (\int T_{44} d\tau_3)_{x_4 = \text{const}} = (\int T_{4l} V_l d\tau_3)_{x_4 = \text{const}}. \end{aligned} \quad (4,5)$$

Die letzte Umformung ist richtig, weil in K gilt: $V_l = V_4 \cdot \delta_{l4}$. Da wir das Integral $-\int T_{44} d\tau_3$ in K als die Ruhenergie $m_0 c^2$ des Teilchens zu deuten haben, so bedeutet die Gleichung (4,5): es ist $(-\int T_{4l} V'_l d\tau'_3)_{x'_4 = \text{const}}$ allgemein die Ruhenergie des Teilchens, wenn wir die Größen T'_{4l} , V'_l durch eine Lorentztransformation $K \rightarrow K'$ aus den für das Ruhesystem K geltenden Größen T_{4l} , $V_l = \delta_{l4}$ berechnen. V'_l ist demnach die Vierergeschwindigkeit (3,5) von K' gegenüber dem Ruhesystem K des Teilchens. Wir können also $-T'_{4l} V'_l$ als Ruhenergie-dichte (in K') auffassen. Daraus ergeben sich sofort die Ausdrücke für die bewegte Energie und den Impuls des Teilchens, welche die richtigen

Transformationseigenschaften haben: Wir brauchen $-T'_{4l} V'_l d\tau'_3$ nur mit $\frac{i}{c} \cdot V'_k$

zu multiplizieren, um Impulsbeitrag und $\frac{i}{c}$ · Energiebeitrag des einzelnen Raumgebietes $d\tau'_3$ in K' zu bekommen. Integration über τ'_3 bei festem x'_4 gibt dann den Impuls-Energie-Vierervektor P'_k ($P_k \equiv \left\{ p, \frac{iE}{c} \right\}$):

$$P'_k \equiv \frac{1}{i c} \left(\int T'_{4l} V_l V'_k d\tau'_3 \right)_{x'_4 = \text{const}}. \quad (4,6)$$

Die Größen sind dabei aus den für das Ruhssystem geltenden durch eine Lorentztransformation $K \rightarrow K'$ zu berechnen wie bei der Ruhenergie (s. oben). Wir rechnen nach, daß die Transformationseigenschaft wirklich die verlangte ist. Herauskommen muß:

$$P'_k \alpha_{\sigma k} = P_{\sigma} = P_4 \delta_{\sigma 4}, \quad (4,7)$$

weil im Ruhssystem nur P_4 von 0 verschieden ist. (Wir machen übrigens darauf aufmerksam, daß unsere Integrale über ein beliebiges Raumstück gehen, und dann zum Beispiel $-\int T_{4l} V_l d\tau_3$ im allgemeinen nur einen Teil der Ruhenergie umfaßt. Die Grenzen der Integrale sind ebenfalls durch Lorentztransformation aus einander zu gewinnen.) Es wird:

$$\begin{aligned} \left(\int T'_{4l} V_l V'_k d\tau'_3 \right)_{x'_4 = \text{const}} &= (T_{nj} \alpha_{n4} \alpha_{jl} V_s \alpha_{sl} V_r \alpha_{rk} \alpha_{\sigma k} d\tau'_3)_{x'_4 = \text{const}} \\ &= \left(\int T_{nj} \alpha_{n4} \delta_{js} V_s V_r \delta_{r\sigma} d\tau'_3 \right)_{x'_4 = \text{const}} = \left(\int T_{ns} \alpha_{n4} V_s^2 \delta_{s4} \delta_{\sigma 4} d\tau'_3 \right)_{x'_4 = \text{const}} \\ &= \left(\int T_{44} \delta_{n4} \alpha_{n4} \delta_{\sigma 4} d\tau'_3 \right)_{x'_4 = \text{const}} = \left(\int T_{44} \delta_{\sigma 4} d\tau_3 \right)_{x_4 = \text{const}} \\ &= \left(\int T_{4l} V_l V_{\sigma} d\tau_3 \right)_{x_4 = \text{const}} = \left(\int T_{4l} V_l V_4 \delta_{\sigma 4} d\tau_3 \right)_{x_4 = \text{const}}, \end{aligned} \quad (4,8)$$

was zu beweisen war. Der Beweis gilt für jedes gestrichene System K' , und damit auch für die Transformation von einem gestrichenen System K' zu einem neuen K'' , weil aus $P'_k \alpha_{\sigma k} = P''_n \gamma_{\sigma n}$ (γ sind die Richtungscosinus der Lorentztransformation $K'' \rightarrow K$) geschlossen werden kann auf: $P'_i = P'_k \epsilon_{ki}$, mit $\epsilon_{ki} = \alpha_{\sigma k} \gamma_{\sigma i}$; die ϵ_{ki} haben alle Eigenschaften der Koeffizienten einer Lorentztransformation (Orthogonalität, Realitätsverhältnisse), wie man leicht bestätigt.

Außer T_{kn} würde jeder Tensor Z_{kn} dieselben Dienste leisten, wenn er die Eigenschaften hat, daß für ein ruhendes Teilchen gilt: $Z_{4n} = \text{const} \cdot T_{44} \cdot \delta_{4n}$. Würden wir nicht T_{kn} als Energie-Spannungstensor deuten, sondern Z_{kn} , so würde schon genügen, daß für ein ruhendes Teilchen: $Z_{4n} = Z_{44} \cdot \delta_{4n}$ wird.

Damit ist die Forderung 6, § 2 erfüllt. Die Forderung 7 von § 2 zeigt, daß die S_l aus dem Momentanzustand des Feldes bestimmt sein sollen; es müssen also außer den Feldgleichungen Beziehungen zwischen den S_l und F_{kn} bestehen, in denen keine Differentialquotienten der F_{kn} vorkommen. Würden nämlich in einem System nur räumliche Differentialquotienten in den Beziehungen vorkommen, so würden wegen der Lorentztransformation in einem anderen System auch zeitliche vorkommen. Andererseits ist S_4 aus den Feldgleichungen durch den Momentanzustand bestimmt, nämlich durch $\text{div } \mathfrak{E}$. Die gesuchten Beziehungen zwischen den S_l und den F_{kn} dürfen also nur erlauben, 3 von den S_l durch das übrige bleibende auszudrücken. Natürlich müssen diese Beziehungen auch Lorentzinvariant sein. Die einfachste, vielleicht einzige Möglichkeit, diese Bedingungen zu erfüllen, ist die, für die S_l vier homogene lineare Gleichungen vorzuschreiben, deren Koeffizienten Funktionen der F_{kn} sind. Wegen der Lorentzinvarianz kommt nur die Form in Frage: $S_n \cdot D_{kn} = 0$, wo die D_{kn} einen Tensor zweiter Stufe bilden, für den man immer die Form: $D_{kn} = H_{kn} - D \cdot \delta_{kn}$ voraussetzen kann. Da für ein ruhendes Teilchen gilt: $S_n = S_4 \cdot \delta_{4n}$, so folgt $S_4 H_{k4} = D S_4 \cdot \delta_{k4}$. Die H_{kn} können demnach die F_{kn} nicht linear enthalten; denn H_{kn} kann nicht

proportional F_{kn} sein, weil $F_{44} = 0$, und weil $D = 0$ allgemein auf $S_k F_{kn} = 0$ führen würde, was wegen $\|F_{kn}\| = -(\mathfrak{E} \mathfrak{H})^2 \neq 0$ unmöglich ist. H_{kn} kann auch keine lineare Funktion der F_{ij} sein, weil man dann als Koeffizientenschema dieses linearen Zusammenhangs einen Tensor vierter Stufe bekäme, dessen Komponenten konstant und physikalisch kaum deutbar wären. H_{kn} muß also die F_{ij} mindestens bilinear enthalten; dies führt notwendig auf $H_{kn} = F_{kj} F_{nj} = T_{kn} + \frac{1}{2} \delta_{kn} F_{ij} F_{ij}$; den zweiten Summanden kann man zu $D \delta_{kn}$ nehmen. Ein allgemeinerer linearer Zusammenhang zwischen H_{kn} und T_{ij} würde wieder physikalisch kaum deutbare Komponenten eines Tensors zweiter oder vierter Stufe bringen. Es bleibt also nur die Form:

$$S_n T_{kn} = B S_k, \quad (4,9)$$

wobei wir gleich die in § 3 eingeführte Bezeichnung B gewählt haben, weil die zum homogenen Gleichungssystem für die S_k gehörige Determinante in der Tat wie dort $\|T_{kn} - B \delta_{kn}\|$ lautet. Für das ruhende Teilchen haben wir $S_k T_{k4} = B S_4 \delta_{k4}$, was für $k \neq 4$ identisch erfüllt ist, und für $k = 4$ liefert: $T_{44} = B$. Da andererseits

nach (3,17): $B = \pm \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}$, was in unserem Fall sich auf $B = \pm W$ vereinfacht, so sehen wir, daß wir wegen $T_{44} = -W$ schließen müssen, daß in unserem Falle nur das untere Vorzeichen gilt: $B = -W$. Allgemein wird also in $B = \pm A$ nur das untere Vorzeichen gelten können, so daß Gleichung (3,17) sich präzisiert zu:

$$B = -\sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} \equiv A, \quad (4,10)$$

eine Behauptung, die wir noch auf anderem Weg bestätigen werden. (4,9) wird damit für ruhende Teilchen eine Identität!

Die Gleichungen (4,9) werden also von jeder Feldverteilung identisch erfüllt, welche für ein ruhendes Teilchen $S_n = S_4 \delta_{n4}$ und $\mathfrak{E} = 0$ liefert. Sie gelten deshalb auch für jedes Teilchen identisch, das im Ruhesystem nur ein elektrostatisches Feld hat, also auch für die Coulombsche Punktladung, und sie gelten nicht nur im Ruhesystem, sondern in beliebigen anderen Systemen, wenn man die Felder für diese Systeme durch Lorentztransformation aus dem Ruhfeld berechnet.

Wird z. B. das Teilchen oder allgemeiner ein beliebiges elektrostatisches Feld gegenüber dem Beobachtersystem K' (Bezeichnung wie in (4,1)) mit der Geschwindigkeit v' in der x -Richtung bewegt, so bestehen in diesem System demnach die merkwürdigen Beziehungen:

$$\varrho' \left(\frac{v'_x}{c} T'_{11} - \frac{\mathfrak{E}'_x}{c} \right) = -\sqrt{W'^2 - \frac{\mathfrak{E}'^2}{c^2}} \cdot \frac{\varrho' v'_x}{c}; \quad \varrho' \left(\frac{v'_x}{c} T'_{12} - \frac{\mathfrak{E}'_y}{c} \right) = 0;$$

$$\varrho' \left(\frac{v'_x}{c} T'_{13} - \frac{\mathfrak{E}'_z}{c} \right) = 0; \quad \varrho' \left(\frac{(v' \mathfrak{E}')}{c^2} - W' \right) = -\sqrt{W'^2 - \frac{\mathfrak{E}'^2}{c^2}} \cdot \varrho'.$$

Setzt man hier den Zusammenhang ein, der zwischen magnetischem und elektrischem Feld für ein bewegtes elektrostatisches Feld besteht:

$$\mathfrak{H}' = \left[\frac{v'}{c} \mathfrak{E}' \right],$$

so kann man die vorher genannten Gleichungen bestätigen, die also unabhängig von der Form von \mathfrak{E}' und für jedes v' gelten. Den einfachsten Beweis gibt natürlich die Transformation aufs Ruhesystem, wo in den ersten 3 Gleichungen alle Glieder Null sind und in der letzten die Identität $-\varrho W = -\varrho' W$ übrig bleibt.

Für die Coulombsche Punktladung ist zu beachten, daß die S_n nur an der Punktsingularität selbst von null verschieden sind, und daß sie dort, wie auch T_{kn} , divergieren. Deshalb wohl hat man die Gleichung (4,9) bisher nicht bemerkt.

Man kann die Gleichungen (4,9) durch die Invariante C dividieren und erhält:

$$V_n T_{kn} = A \cdot V_k. \quad (4,11)$$

Aus den vorigen Bemerkungen geht schon hervor, daß (4,9) nicht ausreichen kann zur Berechnung der S_n . In der Tat haben wir in § 3 gezeigt, daß die Determinante $\|T_{kn} - B \delta_{kn}\|$ den Rang 2 hat; man kann aus (4,9) also nur 2 von den S_n durch die übrigen 2 ausdrücken. Ein weiteres von den S_n , nämlich S_4 , wird durch $\text{div } \mathfrak{E}$ gegeben.

Die noch fehlende Gleichung kann wegen der Kovarianzforderung nur eine Beziehung für die Invariante von S_n sein, also für C . Es ist ja:

$$S_n S_n = C^2 V_n V_n = C^2. \quad (4,12)$$

C muß wegen der Forderung 7 von § 2 Funktion der Invarianten der F_{kn} sein, also von:

$$D = \frac{1}{2} (\mathfrak{H}^2 - \mathfrak{E}^2) = \frac{1}{2} F_{js} F_{js}; \quad K = (\mathfrak{E} \mathfrak{H}); \quad \|F_{kn}\| = -(\mathfrak{E} \mathfrak{H})^2. \quad (4,13)$$

Nach (3,15) ist: $A^2 = D^2 + K^2$.

Wir können also C als von A und $\frac{D}{A}$ oder D abhängig betrachten, und $-A = \zeta$ und $-\frac{D}{A} = \sin \eta$ als Polarkoordinaten ζ, η in einer K, D -Ebene deuten. $A = 0$ bedeutet $D = 0, K = 0$ wegen (3,15), und entspricht dem Nullpunkt in der K, D -Ebene. Weil $C = \frac{S_4}{V_4}$ rein imaginär ist, schreiben wir:

$$iC = f(A, D). \quad (4,14)$$

Wie diese reelle Funktion f aussieht, kann nur die Erfahrung oder die Quantentheorie lehren. iC deuten wir als Koordinate senkrecht zur K, D -Ebene; dann kann (4,14) als Gleichung einer Fläche über der K, D -Ebene aufgefaßt werden. Wenn C nur von A abhängt, ist die Fläche rotationssymmetrisch um die iC -Achse.

Wir werden in § 6 den Ansatz:

$$C = \text{const} \cdot A^n \quad (4,15)$$

für $n \geq 1$ näher untersuchen.

Jedenfalls können wir feststellen, daß mit (4,9) (4,12), (4,14) die Forderung 7 von § 2 erfüllt ist. Aus den Feldgrößen für $t = 0$ kennt man für diesen Zeitpunkt A und D , also nach (4,14) C , wenn die Funktion f bekannt ist, wie es etwa in (4,15) der Fall ist. Dann kann man mit Hilfe von (4,9), (4,12) drei von den S_n durch eines von ihnen ausdrücken, etwa durch S_4 . Wir werden diese Auflösung in § 5 explizit angeben. Aus $\text{div } \mathfrak{E}$ ist auch S_4 bekannt, und damit alle S_n . Die Feldgleichungen — wir übernehmen die Maxwell-Lorentzschen — erlauben dann $\mathfrak{E}, \mathfrak{H}$ zu berechnen im ganzen Raum.

Dies läßt sich durch partielle Differentiation nach t fortsetzen. Da die Gleichungen (4,9), (4,12), (4,14) für jeden Raumzeitpunkt gelten sollen, können sie beliebig differenziert werden, ebenso wie die Feldgleichungen. Man erhält so aus (4,14) \dot{C} und aus (4,12), (4,9) drei von den \dot{S}_n ausgedrückt durch das letzte von den \dot{S}_n . Die Gleichung für $\text{div } \mathfrak{E}$ gibt dann \dot{S}_4 und damit alle \dot{S}_n . Aus den Feldgleichungen folgt so $\dot{\mathfrak{E}}, \dot{\mathfrak{H}}$, und so geht das weiter. Es ergibt sich die ganze Taylorsche Reihe für $t = 0$ im ganzen Raum und damit die völlige Bestimmtheit der

Weiterentwicklung des Feldes, wenn der Zustand \mathcal{E}, \mathcal{H} für $t = 0$ gegeben ist. Dieser Schluß ist unabhängig von der speziellen Form der Funktion f in (4,14). Er zeigt zugleich, daß die Gln. (4,9), (4,12), (4,14) verträglich sind mit den Feldgleichungen.

Die V_n ergeben sich aus den S_n mit Hilfe der Gleichung: $S_n = C \cdot V_n$. Man kennt dann wegen (4,6) auch Impuls und Energie der Teilchen.

Das Gleichungsschema ist fertig, wenn wir noch die Feldgleichungen anschreiben. Wir übernehmen, wie gesagt, die Gleichungen der klassischen Elektronentheorie; die einzige Änderung oder besser: Ergänzung, die wir eingeführt haben, ist der Zusammenhang zwischen S_n und den Feldgrößen, der in (4,9), (4,10), (4,12), (4,13), (4,14) niedergelegt ist. Gegenüber der üblichen Deutung der Elektronentheorie kommt allerdings noch hinzu, daß wir von kontinuierlich verteilter Ladung ρ sprechen, daß wir also die stillschweigende Vereinbarung, die bei der Teilchenbeschreibung in der klassischen Elektrodynamik immer gemacht wird, daß Teilchen punktförmig seien, nicht anerkennen.

Für die partielle Differentiation einer Größe H_k nach der Koordinate x_i verwenden wir die abgekürzte Schreibweise $H_{k,i}$:

$$H_{k,i} \equiv \frac{\partial H_k}{\partial x_i}. \quad (4,16)$$

Dann lauten die Feldgleichungen:

$$F_{kn,n} = S_k; \quad (4,17)$$

$$F_{ki,n} + F_{ln,k} + F_{nk,i} = 0. \quad (4,18)$$

Aus (4,17) folgt wegen der Antisymmetrie von F_{kn} und der Vertauschbarkeit der Differentiationen die Identität:

$$F_{kn,n,k} = 0, \quad (4,19)$$

welche die Erhaltung der Ladung ausspricht:

$$S_{k,k} = 0. \quad (4,19a)$$

(4,18) wird bekanntlich identisch erfüllt durch den Ansatz:

$$F_{ki} = \Phi_{i,k} - \Phi_{k,i}; \quad (4,20)$$

das Viererpotential Φ_i kann immer ersetzt werden durch $\Phi_i + G_i$ mit beliebigem G (Eichinvarianz).

Man kann die Forderung stellen:

$$\Phi_{i,i} = 0; \quad (4,21)$$

dann muß wegen (4,17) auch

$$G_{i,i} = \text{const.} \quad (4,22)$$

verlangt werden; im übrigen bleibt G willkürlich. Wenn (4,21) gefordert wird, erhält man aus (4,17):

$$\Phi_{k,n,n} = -S_k. \quad (4,23)$$

In diesem § haben wir dreierlei aufgestellt: I. die Feldgleichungen (4,17), (4,18); II. die Gleichungen, welche den Viererstrom S_n aus den Feldgrößen zu berechnen erlauben: (4,9) – (4,12), (4,14); III. die Definition für die Ruhmasse (4,5), und für den Impuls-Energie-Vierervektor (4,6). Diese 3 Annahmengruppen sind unabhängig voneinander. Man könnte z. B. an Stelle von III. eine andere Definition wählen, wie wir dort schon erwähnten. Andere Gleichungen als II. zur Erfüllung der Forderung 7, § 1 aufzustellen, dürfte nicht leicht sein. An den Feldgleichungen I. halten wir in dieser Arbeit fest.

Wir untersuchen jetzt die Folgerungen aus II.

§ 5. Die Gleichung $S_n T_{kn} = A S_k$

Wir gehen von (4,9) aus und holen zunächst den versprochenen Beweis nach, daß $B = -\sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}$, also gleich $+A$ gesetzt werden muß, wie wir schon in (4,10) aus dem Spezialfall ruhender Teilchen mit statischen Feldern geschlossen hatten. Wegen (3,17) haben wir zunächst $B = \pm A$, und nach (4,9) für $k = 4$: $S_n T_{4n} = B S_4$. Wir dividieren durch $i\varrho$ und erhalten:

$$\frac{(v \mathfrak{E})}{c^2} - W = \mp \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}. \quad (5,1)$$

Weiter folgt durch Quadrieren:

$$\frac{\mathfrak{E}^2}{c^2} \cdot (\beta^2 \cos^2(v \mathfrak{E}) + 1) = 2W \left(W \mp \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} \right). \quad (5,2)$$

Mit der Abkürzung $\delta = \frac{\mathfrak{E}^2}{c^2 W^2}$ folgt: $\beta^2 \cos^2(v \mathfrak{E}) + 1 = \frac{2(1 \mp \sqrt{1-\delta})}{\delta}$. Nach (3,15) ist $\delta \leq 1$; da δ auch Null sein kann, kommt nur das obere Vorzeichen in Frage, wie in (3,10) behauptet worden war. Also ist allgemein:

$$B = -\sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} = A,$$

und die Gleichungen (4,9) lauten endgültig:

$$S_k T_{kn} = A S_k. \quad (5,3)$$

Division durch C läßt daraus (4,11) entstehen. (5,1) wird jetzt:

$$\frac{(v \mathfrak{E})}{c^2} = W - \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}; \quad (5,4)$$

wegen $W^2 \geq \frac{\mathfrak{E}^2}{c^2}$ ist die Wurzel immer reell, also auch der $\cos(v \mathfrak{E})$. Da die rechte Seite von (5,4) immer ≥ 0 ist, so ist $\cos(v \mathfrak{E})$ immer ≥ 0 . Falls der Winkel $(v \mathfrak{E})$ also reell ist, ist er ein spitzer Winkel. Wenn $v = 0$, folgt: $\delta = 2(1 - \sqrt{1-\delta})$, was auf $\delta = 0$ führt; $v = 0$ ist also nur dort möglich, wo $\mathfrak{E} = 0$ ist. Die Funktion $\frac{2(1 - \sqrt{1-\delta})}{\delta}$ wächst mit steigendem δ ; für $\delta = 0$ hat sie den Wert 1, für $\delta = 1$ den Wert 2; $\beta^2 \cos^2(v \mathfrak{E})$ liegt also zwischen 0 und 1, wie es sein muß. $\delta = 1$ gibt wegen (5,4): $\beta \cos(v \mathfrak{E}) = 1$.

Der Winkel (v, \mathfrak{E}) ist dann reell, wenn $\cos^2(v, \mathfrak{E}) \leq 1$ ist; das bedeutet:

$$\beta^2 \geq \frac{\left(W - \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} \right)^2}{\frac{\mathfrak{E}^2}{c^2}} = \frac{T_{44} - A}{T_{44} + A}. \quad (5,5)$$

Wegen $\beta \leq 1$ (sonst verliert V_k seinen Sinn) folgt aus (5,5) für $\delta = 1$: $\beta = 1$ und $\cos(v \mathfrak{E}) = 1$. Im Fall $A = 0$ gilt also: $|v| = c$; $v \parallel \mathfrak{E}$; $|\mathfrak{E}| = cW$; $|\mathfrak{E}| = |\mathfrak{H}|$, $\mathfrak{E} \perp \mathfrak{H}$.

Mit (5,5) gleichbedeutend ist wegen $\beta \leq 1$:

$$V_4^2 \geq \frac{T_{44} + A}{2A}. \quad (5,6)$$

Die rechte Seite von (5,6) ist ≥ 1 , gleich Eins ist sie nur für $\mathfrak{E} = 0$. Die Bedingung (5,6) werden wir als Bedingung für reelle Geschwindigkeit v wieder-

finden. (5,6) garantiert also erstens, daß die Geschwindigkeit v reell ausfällt, zweitens, daß $\beta \leq 1$ wird, drittens, daß der Winkel zwischen v und \mathfrak{S} reell ist.*

Indices, die nur von 1–3, nicht bis 4 laufen können, kennzeichnen wir durch Unterstreichen. In dieser Bezeichnung wird aus (5,3) für $k \neq 4$ nach Division durch $\frac{\rho}{c}$:

$$\underline{v}_i T_{kj} = \underline{\mathfrak{S}}_k + A v_k. \quad (5,7)$$

Für $k = 4$ entsteht so aus (5,3) die Gl. (5,4), die wir schon behandelt haben. Dabei ist zu beachten, daß die v_k einen dreidimensionalen Vektor, keinen Vierervektor bilden.

Läßt man hier $v \rightarrow c$ gehen, setzt etwa $v_x \rightarrow c$; $v_y = 0 = v_z$, so wird aus der ersten Gl. (5,7): $T_{11} = -i T_{14} + A$, und aus (5,4): $-i T_{14} = -T_{44} + A$, also: $T_{11} - T_{44} = -2i T_{14}$, woraus folgt: $\mathfrak{E}_y = \mathfrak{H}_z$; $\mathfrak{E}_z = -\mathfrak{H}_y$; $\mathfrak{E}_x, \mathfrak{H}_x$ bleiben beliebig; damit sind auch die anderen 2 Gln. (5,7) erfüllt, die lauten: $T_{12} = -i T_{24}$; $T_{13} = -i T_{34}$.

Dies Ergebnis ist ein weiteres wichtiges Argument zugunsten der Gln. (5,3). Man kann nämlich erstens zeigen, daß dieses Ergebnis herauskommen muß, und zweitens, daß nur Gleichungen geeignet erscheinen, es zu liefern, welche den Viererstrom S_k mit den Feldstärken $\mathfrak{E}, \mathfrak{H}$ allein verknüpfen, ohne die Differentialquotienten der Feldstärken zu enthalten. Erstens: Transformiert man einen beliebigen Vierervektor A'_k oder einen beliebigen Tensor 2. Stufe A'_{kl} , der sich auf ein gegen das Beobachtersystem K mit v bewegtes System K' bezieht, von K' auf K , so folgt allein aus der Lorentztransformation, daß für $v_x \rightarrow c$, $v_y = 0 = v_z$, in K gelten muß:

$$i A_1 = A_4, \quad (5,8)$$

für den Vierervektor A_k , und:

$$i A_{12} = A_{42}, i A_{13} = A_{43}, i A_{21} = A_{31}, i A_{31} = A_{21}, i (A_{11} - A_{44}) = A_{14} + A_{41} \quad (5,9)$$

für den Tensor A_{kl} ³⁾. Ist der Tensor antisymmetrisch, so bleiben nur die ersten zwei Beziehungen übrig, die für den Sechservektor des Feldes bedeuten: $\mathfrak{H}_z = \mathfrak{E}_y$; $\mathfrak{H}_y = -\mathfrak{E}_z$. Man erkennt auch, daß die im vorigen Absatz genannten Grenzformen der Gln. (5,7) und (5,4) Spezialfällen der Beziehungen (5,9) entsprechen, wie sie für einen symmetrischen Tensor A_{kl} im limes $v_x \rightarrow c$ gelten müssen. Zweitens: Die Absicht unserer Theorie und jeder einheitlichen Feldtheorie ist, einen Zusammenhang zwischen dem Viererstrom und den Feldstärken anzugeben, wobei wir zunächst noch offen lassen, ob in diesem Zusammenhang auch die Ableitungen der Feldstärken vorkommen können (wir lassen also unsere Forderung 7 von § 2 vorerst unbeachtet, nach der solche Ableitungen nicht vorkommen können). Das verlangt Gleichungen zwischen S_k und den Feldgrößen. Andererseits müssen diese Gleichungen, zusammen mit den Feldgleichungen, alle Aussagen über S_k enthalten. S_k bedeutet, daß an der betrachteten Stelle die Strömung V_k herrscht. Gilt dort $v_x = c$, $v_y = 0 = v_z$, so muß dort auch gelten: $\mathfrak{E}_y = \mathfrak{H}_z$, $\mathfrak{E}_z = -\mathfrak{H}_y$. Diese Aussagen müssen sich also notwendig allein schon aus dem Zusammenhang zwischen dem Viererstrom und den Feldgrößen, und aus den Feldgleichungen ergeben, ohne die Hilfe der Lorentztransformation — eben weil die genannten Gleichungen bereits alle Aussagen über S_k enthalten müssen. Unsere Gleichungen

³⁾ Es ist nicht schwer, das allgemeine Schema anzugeben, nach dem solche Beziehungen für Tensoren beliebiger Stufe ohne Rechnung angeschrieben werden können.

leisten dies wirklich. Man sieht aber auch, daß Beziehungen zwischen Viererstrom und Feldgrößen, welche noch Ableitungen der Feldstärken enthalten würden, die geforderten Gleichungen $\mathcal{E}_y = \mathcal{H}_z$, $\mathcal{E}_z = -\mathcal{H}_y$ nicht als bloße Folge aus dem Limes $v_x \rightarrow c$, $v_y = 0 = v_z$ liefern könnten; es kämen dann auch die Ableitungen der Feldstärken in den Grenzformen der Gleichungen vor.

Im Limes $v_x \rightarrow c$, $v_y = 0 = v_z$ wird $A = -\frac{1}{2}(\mathcal{E}_x^2 + \mathcal{H}_x^2)$; $D = \frac{1}{2}(\mathcal{H}_x^2 - \mathcal{E}_x^2)$. Wo bewegte Ladung vorhanden ist, ist der Ausdruck $iC = -\rho\sqrt{1-\beta^2} = f(A, D)$ auch für $\beta \rightarrow 1$ von Null verschieden, und im allgemeinen raumzeitlich veränderlich, weil er (bis aufs Vorzeichen) die im System K' gemessene Raumladung ρ' vorstellt. Für bewegte Ladung kann nicht gleichzeitig $\mathcal{E}_x = 0$, $\mathcal{H}_x = 0$ sein für $\beta \rightarrow 1$, weil sonst die veränderliche Größe iC gleich einer Konstanten, nämlich gleich $f(0, 0)$ sein müßte. In der Tat ist z. B. für das in § 6 untersuchte Modell eines Teilchens, aber auch schon für die bewegte Coulombsche Ladung, für $\beta \rightarrow 1$ zwar $\mathcal{H}_x = 0$, aber $\mathcal{E}_x = \mathcal{E}'_x$ von Null verschieden.

Dagegen ist im ladungsfreien Gebiet $iC = f(A, D) = 0$ für alle β . Im Grenzfall $v_x \rightarrow c$, $v_y = 0 = v_z$ bedeutet dies, daß \mathcal{H}_x Funktion von \mathcal{E}_x wird. Die Integration der Feldgleichungen ist für diesen Fall einfach und zeigt, daß \mathcal{E}_x , \mathcal{H}_x beide raumzeitlich konstant sein müssen. Für ungeladene, mit Lichtgeschwindigkeit wandernde Teilchen ergibt sich als zugehöriges wanderndes Feld die transversale Welle. Die Voraussetzung, daß \mathcal{E} , \mathcal{H} nur von x, t abhängen, die sonst beim Beweis der Transversalität gemacht wird, haben wir nicht benutzt. Die Welle wandert in der x -Richtung mit der Phasengeschwindigkeit c ; es ist:

$$\mathcal{E}_x = 0; \quad \mathcal{H}_x = 0;$$

$$\mathcal{E}_y = \mathcal{H}_z = \text{Im } \varphi(x - ct, y + iz); \quad \mathcal{E}_z = -\mathcal{H}_y = \text{Re } \varphi(x - ct, y + iz);$$

φ ist eine willkürliche komplexe Funktion. Diese Lösung genügt natürlich auch den Maxwellschen Vakuumgleichungen, für $\rho = 0$ sind unsere Feldgleichungen ja mit den Maxwellschen identisch. Neu ist, daß wir aus $v_x \rightarrow c$, $v_y = 0 = v_z$ und aus der Ladungsfreiheit auf transversale Welle schließen können. Es gilt: $A = 0 = D$; $|\mathcal{E}| = |\mathcal{H}|$; $\mathcal{E} \perp \mathcal{H}$; $vW = \mathcal{E}$; $|v| = c$. Die Teilchen wandern in Richtung der Welle.

Für $C = 0$ und beliebiges β braucht keineswegs auch $A = 0$ zu sein; es besteht dann wegen $f(A, D) = 0$ ein Zusammenhang zwischen $\mathcal{H}^2 - \mathcal{E}^2$ und $(\mathcal{E}\mathcal{H})$. Für $A \neq 0$ hätte man damit Lösungen für ungeladene Teilchen endlicher Ruhmasse, wenn das Integral $\int A V_4 d\tau_3$ konvergiert. Nach § 4 ist ja $\int T_{41} V_1 d\tau_3 = \int A V_4 d\tau_3$ proportional der Ruhenergie.

Wenn umgekehrt $A = 0$, so folgt: $|\mathcal{E}| = |\mathcal{H}|$; $\mathcal{E} \perp \mathcal{H}$; $vW = \mathcal{E}$; $|v| = c$; $D = 0$. Liegt v in der x -Richtung, so wird $\mathcal{E}_x = 0$, $\mathcal{H}_x = 0$, was bewegte Ladung ausschließt. Das bedeutet: $iC = f(0, 0) = \text{const} = -\rho'$; bewegte Ladung kann damit nicht dargestellt werden. Es muß $f(0, 0) = 0$ sein, damit der Zusammenhang $iC = f(A, D)$ wirklich einer Theorie von Teilchen entsprechen kann. $f = 0$ gibt $\rho' = 0$: im System K' ist keine Ladung vorhanden (in dem Gebiet, in welchem $f = 0$); weiter ist $\rho = 0$.

Auch über das Verhalten von f in der Nähe von $A = 0$ kann man Aussagen machen. $A = 0$ kann auch durch $\mathcal{E} = 0$, $\mathcal{H} = 0$ realisiert werden; lassen wir \mathcal{E} , \mathcal{H} von Null an wachsen, so wächst A und f auch von Null an. f wird in diesem Wertebereich nicht beliebig schnell mit A wachsen, weil ρ nicht beliebig schnell wachsen wird. Also muß in der Nähe von $A = 0$ entweder sein:

$$iC \sim A^n; \quad n \geq 1 \quad (5,10)$$

oder: $A = 0$ ist ein wesentlich singulärer Punkt von $f(A, D)$, wo f stärker Null wird als jede endliche Potenz A^n von A . Jedenfalls ist $\varrho = 0$ notwendige Folge von $A = 0$. Dagegen kann $\varrho = 0$ sein, ohne daß $A = 0$ wäre.

Wenn $AV_4 \rightarrow 0$ für $A \rightarrow 0$, ist die Ruhmasse dieser mit Lichtgeschwindigkeit wandernden ungeladenen Teilchen Null. Wir begründen dies, wie es in der gewöhnlichen Theorie der Lichtquanten geschieht: durch Grenzübergang $\beta \rightarrow 1$ aus der Impuls-Energiebeziehung materieller Teilchen.

Nach (4,5) gilt für ein materielles Teilchen — wobei wir allerdings die erwähnte Einschränkung machen müssen, daß es im Ruhssystem ein rein elektrisches Feld hat: $-T_{41} V_1 d\tau_3 = -AV_4 d\tau_3$ ist der Beitrag $dm_0 c^2$ von $d\tau_3$ zu seiner Ruhenergie $m_0 c^2$:

$$-AV_4 d\tau_3 = dm_0 c^2. \quad (5,11a)$$

Und (4,6) zeigt, daß:

$$\frac{1}{i c} A V_4 V_k d\tau_3 = dP_k = \frac{dm_0 \{v_k, i c\}}{\sqrt{1-\beta^2}} = \{dp_k, i dm c\} = dm \{v_k, i c\}, \quad (5,11b)$$

der Beitrag von $d\tau_3$ zum Impulsvierervektor P_k des materiellen Teilchens ist. Es gilt:

$$dP_k V_k = \frac{1}{i c} A V_4 V_k V_k d\tau_3 = \frac{1}{i c} A V_4 d\tau_3 = i dm_0 c. \quad (5,12a)$$

Dies bedeutet wie in der gewöhnlichen Mechanik:

$$\frac{1}{\sqrt{1-\beta^2}} \left(-\frac{i}{c} dp_k v_k + i dm c \right) = \frac{i dm c}{\sqrt{1-\beta^2}} (1-\beta^2) = i dm_0 c. \quad (5,12b)$$

In der üblichen Lichtquantentheorie kann man Lichtquanten wie Materieteilchen der Ruhmasse Null, aber endlicher bewegter Masse (wegen $\beta = 1$), behandeln, soweit man sich nur für ihre Impuls- und Energieeigenschaften interessiert. Die formalen Schlüsse, die in diesem Zusammenhang in der üblichen Theorie gezogen werden, ergeben sich auch in unserer Theorie. Wir übertragen dementsprechend die Aussagen (5,11a) und (5,12b), die nach unserer Theorie für materielle Teilchen gelten, formal auf Teilchen der Ruhmasse Null. (5,11a) lehrt, daß dann $AV_4 \rightarrow 0$ gehen muß, und wegen $V_4 \geq 1$ folgt auch $A \rightarrow 0$. Umgekehrt folgt aus $AV_4 \rightarrow 0$ die Ruhmasse null, wie wir behauptet hatten. Für solche Teilchen gibt (5,12b):

$$(dp, v) = dm c^2 = dE,$$

wenn E die Energie des Teilchens ist. Wegen $dp \parallel v$ bedeutet dies die Gleichung:

$$|dp| = \frac{dE}{c}, \quad (5,13)$$

wie es sich für Lichtquanten gehört. Wäre AV_4 endlich für $A \rightarrow 0$, so würde dP_k unendlich. Wir können also zusammenfassen: Die fortschreitende elektromagnetische Welle im ladungsfreien Raum ist eine Lösung unserer Gleichungen, sie entspricht $A = 0$, $AV_4 = 0$, $\varrho = 0$. In ihr bewegen sich ungeladene „Teilchen“ mit Lichtgeschwindigkeit in Richtung der Wellennormale; sie haben die Ruhmasse null, und ihr Impuls ist gleich ihrer Energie E geteilt durch c . Mit anderen Worten: In unseren nichtquantentheoretischen Gleichungen stecken alle Aussagen über Lichtquanten, die ohne das Plancksche Wirkungsquantum formuliert werden können. Die Behauptung, daß $E = h \nu$ sei, ist natürlich nicht in unserer Theorie enthalten.

Diese Ergebnisse sprechen für den Ansatz (5,3) und zwar haben wir ihn zuletzt in der Form (5,4), (5,7) benutzt.

Wir geben jetzt die Auflösung von (5,3) nach den S_n und V_n . Da die Determinante dieser Gleichungen den Rang 2 hat, kann man 2 von den V_n durch die übrigen 2 ausdrücken. Außerdem besteht die Gleichung $V_n V_n = 1$, so daß man 3 von den V_n durch das vierte darstellen kann.

Wenn wir diese letzte Bedingung zunächst noch nicht beachten, erhalten wir ⁴⁾ mit Hilfe der Sätze von § 3:

$$\left. \begin{aligned} V_1 &= \alpha_{13} V_3 + \alpha_{14} V_4; & V_2 &= \alpha_{23} V_3 + \alpha_{24} V_4; \\ \alpha_{13} &= \frac{1}{\Delta_1} [T_{13} (T_{44} + A) - T_{14} T_{34}]; & \alpha_{14} &= \frac{1}{\Delta_1} [T_{14} (T_{33} + A) - T_{13} T_{34}]; \\ \alpha_{23} &= \frac{1}{\Delta_1} [T_{23} (T_{44} + A) - T_{24} T_{34}]; & \alpha_{24} &= \frac{1}{\Delta_1} [T_{24} (T_{33} + A) - T_{23} T_{34}]; \\ \Delta_1 &= (T_{11} - A) (T_{22} - A) - T_{12}^2 = (T_{33} + A) (T_{44} + A) - T_{34}^2. \end{aligned} \right\} \quad (5,14)$$

Dazu kommt die Bedingung: $V_n V_n = (\alpha_{13}^2 + \alpha_{23}^2 + 1) V_3^2 + (\alpha_{14}^2 + \alpha_{24}^2 + 1) V_4^2 + 2 V_3 V_4 (\alpha_{13} \alpha_{14} + \alpha_{23} \alpha_{24}) = 1$.

Es ist wegen Satz IX), § 3:

$$\alpha_{13}^2 + \alpha_{23}^2 + 1 = \frac{2A}{\Delta_1} (T_{44} + A); \quad \alpha_{14}^2 + \alpha_{24}^2 + 1 = \frac{2A}{\Delta_1} (T_{33} + A); \quad (5,15)$$

$$\alpha_{13} \alpha_{14} + \alpha_{23} \alpha_{24} = -\frac{2A}{\Delta_1} T_{34}.$$

Das gibt:

$$V_3 = \frac{1}{T_{44} + A} \left[T_{34} V_4 \pm \sqrt{\Delta_1 \left(\frac{T_{44} + A}{2A} - V_4^2 \right)} \right]. \quad (5,16)$$

Damit die Geschwindigkeit reell wird, muß V_3 rein imaginär ausfallen. Dazu gibt es noch 2 Möglichkeiten:

$$a) \Delta_1 \geq 0; \quad \frac{T_{44} + A}{2A} \leq V_4^2;$$

$$b) \Delta_1 \leq 0; \quad \frac{T_{44} + A}{2A} \geq V_4^2.$$

Man kann zeigen, daß immer $\Delta_1 \geq 0$, daß also nur der Fall a) die Bedingung für reelle Geschwindigkeit darstellt. $V_4^2 \geq \frac{T_{44} + A}{2A}$ war zugleich die Bedingung dafür, daß der Winkel (v , \mathfrak{E}) reell wird, siehe Gleichung (5,6). Es ist wegen (3,11) und (3,14)

$$\begin{aligned} \Delta_1 &= -W(W - \text{III}) + A(-W + W - \text{III}) + W^2 - \frac{\mathfrak{E}^2 - \mathfrak{E}_3^2}{c^2} \\ &= \text{III} \left(W + \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} \right) - \frac{\mathfrak{E}^2 - \mathfrak{E}_3^2}{c^2}. \end{aligned} \quad (5,17)$$

$\Delta_1 \geq 0$ verlangt:

$$\begin{aligned} &(\mathfrak{E}_2^2 + \mathfrak{E}_3^2) (\mathfrak{E}^2 + \mathfrak{E}_3^2 + \sqrt{(\mathfrak{E}^2 - \mathfrak{E}_3^2)^2 + 4(\mathfrak{E} \mathfrak{E}_3)^2}) \\ &\geq 2 [(\mathfrak{E}_2 \mathfrak{E}_3 - \mathfrak{E}_3 \mathfrak{E}_2)^2 + (\mathfrak{E}_3 \mathfrak{E}_1 - \mathfrak{E}_1 \mathfrak{E}_3)^2], \end{aligned}$$

wobei das Fragezeichen darauf hindeuten soll, daß wir diese Beziehung erst beweisen wollen. Es folgt:

$$(\mathfrak{E}_2^2 + \mathfrak{E}_3^2)^2 \sqrt{(\mathfrak{E}^2 - \mathfrak{E}_3^2)^2 + 4(\mathfrak{E} \mathfrak{E}_3)^2} \geq (\mathfrak{E}^2 - \mathfrak{E}_3^2) (\mathfrak{E}_3^2 - \mathfrak{E}_2^2) - 4 \mathfrak{E}_2 \mathfrak{E}_3 (\mathfrak{E} \mathfrak{E}_3);$$

⁴⁾ Die α der Gl. (5,14) sind nicht etwa Koeffizienten einer orthogonalen Transformation, sondern bloße Abkürzungen.

die linke Seite ist ≥ 0 , weil die Wurzel positiv gemeint ist. Wenn die rechte Seite ≤ 0 ist, so ist die Bedingung von selbst erfüllt. Wir brauchen also nur zu beweisen, daß die Ungleichung auch dann noch gilt, wenn die rechte Seite ≥ 0 ist. Dann kann man beide Seiten quadrieren und erhält schließlich

$$\{(\mathfrak{H}^2 - \mathfrak{E}^2) \mathfrak{E}_3 \mathfrak{H}_3 - (\mathfrak{H}_3^2 - \mathfrak{E}_3^2) (\mathfrak{E} \mathfrak{H})\}^2 \geq 0, \quad (5,18)$$

und dies ist immer erfüllt. Also ist immer $\Delta_1 \geq 0$; die Bedingung für reelle Geschwindigkeit, und für reelle Winkel zwischen v und \mathfrak{E} lautet demnach

$$V_4 \geq \frac{T_{44} + A}{2A} = \frac{W + \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}}{2 \cdot \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}}. \quad (5,19)$$

Der Ausdruck rechts ist immer ≥ 1 , wie es sich für V_4 gehört. Gleichbedeutend mit (5,19) ist:

$$\beta^2 \geq \frac{T_{44} - A}{T_{44} + A} = \frac{W - \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}}{W + \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}}. \quad (5,20)$$

Der Ausdruck rechts ist immer ≤ 1 und ≥ 0 , wie es sein muß. V_4 ist andererseits bestimmt durch:

$$V_4 = \frac{S_4}{C} = \frac{\varrho}{-iC} = -\frac{\text{div } \mathfrak{E}}{f(A, D)}. \quad (5,21)$$

f ist reell; (5,19) verlangt:

$$\frac{(\text{div } \mathfrak{E})^2}{f^2(A, D)} \geq \frac{T_{44} + A}{2A}; \quad (5,22)$$

nur dort, wo (5,22) erfüllt ist, gibt es reelle Geschwindigkeit. Da die T_{ii} rein imaginär sind für $i \neq 4$, T_{44} und A aber immer reell, so wird V_3 rein imaginär, wenn (5,19) erfüllt ist. α_{13}, α_{23} sind reell, α_{14}, α_{24} rein imaginär, daher fallen mit V_3 zugleich auch V_1, V_2 rein imaginär aus.

Durch Einsetzen von (5,16) in (5,14) ergeben sich auch V_1, V_2 in Abhängigkeit von V_4 . Wir führen die Abkürzungen ein:

$$\alpha_{13} = \frac{1}{\Delta_1} [(T_{13} + A \delta_{13}) (T_{44} + A) - (T_{14} + A \delta_{14}) T_{34}]; \quad (5,23)$$

$$\alpha_{14} = \frac{1}{\Delta_1} [(T_{14} + A \delta_{14}) (T_{33} + A) - (T_{13} + A \delta_{13}) T_{34}].$$

Für $l = 1, 2$ gehen sie in die Ausdrücke über, die in (5,14) angegeben sind. Es ist: $\alpha_{33} = \alpha_{44} = 1$; $\alpha_{34} = \alpha_{43} = 0$, und es gilt:

$$\alpha_{13} T_{34} + \alpha_{14} (T_{44} + A) = T_{14} + A \delta_{14}. \quad (5,24)$$

Damit bekommt die Auflösung von (4,11) nach den V_n die Form:

$$V_l = \frac{1}{T_{44} + A} \left[(T_{14} + A \delta_{14}) V_4 \pm \alpha_{13} \cdot i \sqrt{\Delta_1 \left(V_4^2 - \frac{T_{44} + A}{2A} \right)} \right]; \quad l = 1, 2, 3, 4. \quad (5,25)$$

Für $l = 4$ wird (5,25) eine Identität, wie es sein muß; eines von den V_p in unserem Fall V_4 , bleibt ja unbestimmt, und wird erst durch die Anfangsbedingungen ($\mathfrak{E}, \mathfrak{H}$ für $t = 0$ gegeben) festgelegt. Für $l = 3$ geht (5,25) in (5,16) über. Durch Einsetzen bestätigt man, daß die Lösung (5,25) wirklich die Gleichungen: $(T_{ki} - A \delta_{ki}) V_i = 0$ und $V_i V_i = 1$ erfüllt.

Die allgemeine Auflösung dieser Bedingungsgleichungen für die V_i lautet, wenn wir festsetzen, über eingeklammerte Indizes nicht zu summieren:

$$V_n = \frac{1}{T_{(j)} + A} \left[(T_{(j)} + A \delta_{n(j)}) V_{(j)} \pm \alpha_{n(l)}^{(j)} \cdot i \sqrt{\Delta_{(l)}^{(j)} \left(V_j^2 - \frac{T_{(j)} + A}{2A} \right)} \right];$$

$$i \equiv \sqrt{-1}; \quad \alpha_{nl}^{(j)} = \frac{1}{\Delta_{(l)}^{(j)}} \cdot [(T_{n(l)} + A \delta_{n(l)}) (T_{(j)} + A) - (T_{n(j)} + A \delta_{n(j)}) T_{(l)}];$$

$$\Delta_l^j = (T_{(j)} + A) \cdot (T_{(l)} + A) - T_{lj}^2; \quad l \neq j. \quad (5,26)$$

Der Index l darf nur nicht gleich j gewählt werden, im übrigen ist l eine beliebige der Zahlen 1, 2, 3, 4; auch j ist eine beliebige dieser 4 Zahlen. Für $n = j$ wird (5,26) eine Identität. Es gilt:

$$\Delta_l^j = \Delta_j^l; \quad \alpha_{jl}^{(j)} = 0; \quad \alpha_{ll}^{(j)} = 1; \quad (l \neq j). \quad (5,27)$$

(5,25) ist ein Spezialfall von (5,26).

Die Gleichungen $V_n (T_{nk} - A \delta_{nk}) = 0$; $V_n V_n = 1$ sind auch durch (5,27) identisch erfüllt. Auf dieselbe Weise wie oben kann man zeigen, daß: $\Delta_l^j \geq 0$, wenn j oder l gleich 4 ist, und daß in allen anderen Fällen $\Delta_l^j \leq 0$ ist.

Die Bedingungen für reelle Bewegung sind:

$$V_j^2 \leq \frac{T_{(j)} + A}{2A} \quad \text{für } j \neq 4; \quad V_4^2 \geq \frac{T_{44} + A}{2A}. \quad (5,28)$$

Setzt man $v_j \equiv c \beta_j$, so hat man die Bedingungen:

$$\frac{\beta_j^2}{1 - \beta^2} \geq \frac{T_{(j)} - \sqrt{W^2 - \frac{c^2}{c^2}}}{2 \sqrt{W^2 - \frac{c^2}{c^2}}} \quad \text{für } j \neq 4. \quad (5,29)$$

Daraus ergibt sich durch Addition:

$$1 - \beta^2 \leq \frac{2 \sqrt{W^2 - \frac{c^2}{c^2}}}{W - \sqrt{W^2 - \frac{c^2}{c^2}}}; \quad V_4^2 \geq \frac{T_{44} - A}{2A} = \frac{W - \sqrt{W^2 - \frac{c^2}{c^2}}}{2 \sqrt{W^2 - \frac{c^2}{c^2}}}. \quad (5,30)$$

Die Bedingungen für die β_j^2 sind also weniger einschränkend als die Bedingung (5,28) für V_4^2 . In der Tat können die Ausdrücke $T_{(j)} + A < 0$ sein für $j \neq 4$, sie sind es für ein ruhendes Teilchen. In solchen Fällen bedeutet (5,28) gar keine Bedingung für die β_j^2 , wohl aber eine für V_4^2 . Von selbst gilt nach der ersten Gl. (5,30):

$$\frac{T_{(j)} + A}{2|A|} (1 - \beta^2) \leq 1 \quad \text{für } j \neq 4,$$

wegen $\frac{T_{(j)} + A}{-T_{44} + A} \leq 1$. Die Grenze (5,28) für die β_j^2 liegt demnach immer unterhalb 1, höchstens bei 1, während V_4^2 nach (5,28) immer mindestens gleich 1 ist. Die Bedingungen für reelle Bewegung sind also untereinander grundsätzlich verträglich.

Man kann statt von den Gleichungen für die V_n von den Gleichungen (5,3) für die S_n ausgehen; dann steht in (5,26) S_n an Stelle von V_n , und der Ausdruck $\frac{T_{(j)} + A}{2A}$ unter der Wurzel bekommt den Faktor $C^2 \equiv S_i S_i = c^2 (\beta^2 - 1)$. Da die S_j für $j \neq 4$ reell sind für reelle Bewegung, lauten die (5,28) entsprechenden Bedingungen für reelle Bewegung jetzt:

$$S_j^2 \geq \frac{T_{(j)} + A}{2A} \cdot C^2 \quad \text{für } j \neq 4; \quad S_4^2 \leq \frac{T_{44} + A}{2A} \cdot C^2. \quad (5,31)$$

Wenn $\mathfrak{S} = 0$, gilt im allgemeinen $\mathfrak{S} \parallel \mathfrak{E}$, und es wird aus (5,25):

$$V_k = \pm \frac{\mathfrak{E}_k}{|\mathfrak{E}|} \sqrt{1 - V_4^2}; \quad k \neq 4. \quad (5,32)$$

Das bedeutet:

$$v_k = \pm \frac{\mathfrak{E}_k}{|\mathfrak{E}|} |v|. \quad (5,33)$$

Dann ist also $v \parallel \mathfrak{E}$.

Allgemein gilt:

C ist (grundsätzlich) durch (4,14) bestimmt. Mit (5,26) hat man wegen $S_n = C V_n$ alle S_n in Abhängigkeit von den Feldstärkenwerten, die im betrachteten Raumzeitpunkt gelten, und in Abhängigkeit von V_4 , oder von S_4 . Wenn $\text{div } \mathfrak{E}$ bekannt ist, hat man S_4 und damit alle S_k . In § 4 war gezeigt worden, daß aus der Kenntnis von \mathfrak{E} , \mathfrak{S} in einem bestimmten Zeitpunkt das spätere Verhalten von \mathfrak{E} , \mathfrak{S} und S_n völlig bestimmt ist.

Man kann die Gleichungen (4,11) umschreiben in die Form:

$$\begin{aligned} V_n \{F_{ki} F_{ni} - \delta_{kn} A'\} &= 0; \quad A' = A + \frac{1}{4} F_{js} F_{js} = A + D \\ &= \frac{1}{2} \{ \mathfrak{S}^2 - \mathfrak{E}^2 - \sqrt{(\mathfrak{S}^2 - \mathfrak{E}^2)^2 + 4 (\mathfrak{E} \mathfrak{S})^2} \}; \end{aligned} \quad (5,34)$$

A' ist, wie A , immer ≤ 0 .

§ 6. Die Lösung für ein ruhendes Teilchen

Ein ruhendes Teilchen werden wir durch $v = 0$ zu beschreiben haben, was in unserer Theorie bedeutet, daß die Raumladung überall in Ruhe ist. Es gilt dann: $V_k = V_4 \cdot \delta_{k4}$; $V_4 = 1$. Unter (5,4) folgt $\mathfrak{S} = 0$. Wir werden $\mathfrak{S} = 0$ annehmen, also ein rein elektrisches Feld, für das zu Anfang $\text{rot } \mathfrak{E} = 0$ gelten soll, d. h. ein elektrostatisches Feld. Die Theorie muß die Aussage liefern, daß ein solcher Anfangszustand, wenn er einmal bestand, sich beliebig lange hält, wenn keine äußeren Einwirkungen stattfinden, daß das Teilchen also in Ruhe bleibt. Zugleich ist in dieser Aussage das Trägheitsgesetz enthalten. Denn durch eine Lorentztransformation wird aus der vorigen Aussage die Behauptung, daß ein gradlinig gleichförmig bewegtes Teilchen ohne äußere Einwirkung seine Geschwindigkeit nicht ändert. Die Aussage ist also keineswegs trivial, und auch mathematisch wegen des komplizierten Baues der Feldgleichungen nicht selbstverständlich. Es muß sich ja ergeben, daß \mathfrak{E} , \mathfrak{S} , v ständig ihre Anfangswerte beibehalten. Der Beweis läßt sich führen, ohne daß man die Funktion C in ihrer Abhängigkeit von A und D zu kennen braucht, vergleiche (4,14).

Für $t = 0$ haben wir nach Voraussetzung: $\text{rot } \mathfrak{E} = 0$; $\mathfrak{S} = 0$, daher $\dot{\mathfrak{S}} = 0$ und: $\frac{\dot{\mathfrak{E}}_k}{c} = -S_k = -C V_k = 0$; das Unterstreichen bedeutet, wie bisher $k \neq 4$. Es ist also $\dot{\mathfrak{E}} = 0$, $\dot{\mathfrak{S}} = 0$. Außerdem ist: $\text{div } \mathfrak{E} = -i S_4 = -i C V_4 = -i C$. Differenziert man die Feldgleichungen nach t , so kommt: $\text{rot } \dot{\mathfrak{E}} = -\frac{\dot{\mathfrak{S}}}{c}$, woraus $\dot{\mathfrak{S}} = 0$ folgt. Weiter: $\text{rot}_k \dot{\mathfrak{S}} - \frac{\dot{\mathfrak{E}}_k}{c} = \dot{S}_k$; $\text{div } \dot{\mathfrak{E}} = -i \dot{S}_4$. Daraus folgt: $\dot{S}_4 = 0$ und zunächst: $\dot{S}_k = -\frac{\dot{\mathfrak{E}}_k}{c}$. Es ist: $\dot{S}_k = \dot{C} V_k + C \dot{V}_k$; $\dot{C} = f_A \cdot \dot{A} + f_D \cdot \dot{D}$; \dot{A} ist eine lineare Funktion von $\dot{\mathfrak{E}}$, $\dot{\mathfrak{S}}$ und deshalb null für $t = 0$; dasselbe gilt für \dot{D} ,

so daß \dot{C} null wird. Wegen $\dot{S}_4 = 0$ ist demnach $\dot{V}_4 = 0$. V_k ist nach (5,25) Funktion von \mathfrak{E} , \mathfrak{H} , V_4 ; also ist \dot{V}_k eine lineare Funktion von $\dot{\mathfrak{E}}$, $\dot{\mathfrak{H}}$, \dot{V}_4 und damit auch null für $t = 0$. Das bedeutet $\dot{S}_k = 0$ und deshalb $\dot{\mathfrak{E}}_k = 0$.

Man kann leicht zeigen, daß alle zeitlichen partiellen Differentialquotienten der F_{ki} und der S_k null werden. Zum Beweis bezeichnen wir den n ten partiellen zeitlichen Differentialquotienten mit einem eingeklammerten oberen Index n . Wir nehmen an, es sei bewiesen, daß $F_{ki}^{(\alpha)} = 0$ und $S_k^{(\alpha)} = 0$ für $1 \leq \alpha \leq n$; für $n = 1$ trifft dies zu. Wir wollen zeigen, daß daraus folgt, daß $F_{ki}^{(n+1)} = 0$; $S_k^{(n+1)} = 0$.

Aus den Feldgleichungen folgt zunächst: $\mathfrak{H}^{(n+1)} = 0$; $\frac{\mathfrak{E}_k^{(n+1)}}{c} = -S_k^{(n)} = 0$; also $F_{ki}^{(n+1)} = 0$; und weiter: $\text{div } \mathfrak{E}^{(n+1)} = 0 = -i S_4^{(n+1)}$. Wenn wir $\dot{C} = f_A \cdot \dot{A} + f_D \cdot \dot{D}$ τ mal differenzieren, wobei $\tau \leq n$ sein soll, bekommen wir:

$$C^{(\tau+1)} = \sum_{\sigma=0}^{\tau} \{ (f_A)^{(\tau-\sigma)} \cdot A^{(\sigma+1)} + (f_D)^{(\tau-\sigma)} \cdot D^{(\sigma+1)} \} \cdot \binom{\tau}{\sigma}.$$

Ebenso ist:

$$\dot{A} = A_{F_{k1}} \cdot \dot{F}_{k1}; \quad A^{(\sigma+1)} = \sum_{\lambda=0}^{\sigma} \binom{\sigma}{\lambda} (A_{F_{k1}})^{(\sigma-\lambda)} F_{k1}^{(\lambda+1)},$$

also ist $A^{(\sigma+1)} = 0$, weil σ zwischen 0 und n liegt. Dasselbe gilt für $D^{(\sigma+1)}$ aus dem gleichen Grund. Deshalb ist $C^{(\tau+1)} = 0$ für $0 \leq \tau \leq n$. Es ist nach (5,25), wenn man mit C multipliziert denkt:

$$S_k = H_k(F_{1j}, S_4, C);$$

also:

$$S_k^{(n+1)} = \sum_{\sigma=0}^n \{ (H_{k,F_{1j}})^{(n-\sigma)} \cdot F_{1j}^{(\sigma+1)} + (H_{k,S_4})^{(n-\sigma)} \cdot S_4^{(\sigma+1)} + (H_{k,C})^{(n-\sigma)} \cdot C^{(\sigma+1)} \} \cdot \binom{n}{\sigma} = 0,$$

was zu beweisen war.

Solange also keine Einwirkung von außen stattfindet, welche durch die Randbedingungen von Null verschiedene partielle zeitliche Differentialquotienten erzeugen kann, so lange behalten \mathfrak{E} , \mathfrak{H} , v ihre Anfangswerte bei; es bleibt also $\mathfrak{H} = 0$, $v = 0$ und das elektrische Feld \mathfrak{E} genügt dauernd den Gleichungen: $\text{rot } \mathfrak{E} = 0$; $\text{div } \mathfrak{E} = \rho = -i S_4 = -i C = -f(A, D)$. Wegen $D = \frac{1}{2}(\mathfrak{H}^2 - \mathfrak{E}^2)$ und $\mathfrak{E} = 0$ läuft dies darauf hinaus, daß $\text{div } \mathfrak{E}$ Funktion von \mathfrak{E}^2 ist. Aus (4,6) folgt, daß der Impuls des Teilchens null ist, wegen $V_k = 0$, während seine Energie E sich bestimmt aus: $P_4 = \frac{iE}{c} = \frac{1}{ic} (\int T_{44} d\tau_3)_{x_4 = \text{const}}$. Das bedeutet:

$$E = - \int T_{44} d\tau_3 = \int W d\tau_3 = E_0 = m_0 c^2, \quad (6,1)$$

denn dieses E ist Ruhenergie, weil $-(\int T_{4i} V_i d\tau_3)_{x_4 = \text{const}}$ nach § 4 die Ruhenergie vorstellt.

Wir wollen jetzt mit dem Ansatz (4,15) das Problem des ruhenden Teilchens zunächst für $n = 1$ rechnen. Wir setzen:

$$iC = 2\gamma A, \quad (6,2)$$

wo γ eine Konstante ist, die aus Dimensionsgründen hinzugefügt werden muß; γ ist reell, denn $C = \frac{S_4}{V_4}$ ist rein imaginär. Wegen $\mathfrak{H} = 0$ erhalten wir für unseren

Fall $C = i\gamma \mathfrak{E}^2$. Für das skalare Potential φ des Teilchenfeldes entsteht die Differentialgleichung:

$$\Delta\varphi = -\gamma (\text{grad } \varphi)^2; \quad (6,3)$$

und es ist: $\mathfrak{E} = -\text{grad } \varphi$. Die Gleichung (6,3) läßt sich auf die Potentialgleichung zurückführen.

Wir zeigen dies für den allgemeineren Typ der Differentialgleichung

$$\Delta u = G(u) \cdot (\text{grad } u)^2, \quad (6,4)$$

wo G eine beliebig gegebene Funktion von u sein kann. Macht man nämlich in der Potentialgleichung $\Delta\Phi = 0$ die Substitution:

$$\Phi = g(u), \quad (6,5)$$

so entsteht: $g_u \Delta u + g_{uu} \cdot (\text{grad } u)^2 = 0$. Setzen wir also: $G(u) = -\frac{g_{uu}}{g_u}$, so entsteht (6,4). Dies gibt:

$$g(u) = \int^u e^{-\int G(\sigma) d\sigma} d\sigma. \quad (6,6)$$

g kann berechnet werden, weil G nach Voraussetzung bekannt ist. Die Umkehrfunktion zu g nennen wir g^{-1} , und bekommen aus (6,5):

$$u = g^{-1}(\Phi); \quad (6,7)$$

dies ist die allgemeine Lösung der Gleichung (6,4), wenn Φ die allgemeine Lösung der Potentialgleichung ist. Der Beweis gilt für beliebig viel Dimensionen, dann hat man eben unter Δu und $(\text{grad } u)^2$ die entsprechenden n -dimensionalen Ausdrücke zu verstehen. Für 2 Dimensionen erhält man so den bemerkenswerten Satz, daß die allgemeine Lösung von (6,4) durch (6,6), (6,7) gegeben ist, wo $\Phi = f_1(x + iy) + f_2(x - iy)$ aus 2 willkürlichen Funktionen f_1, f_2 der angeschriebenen komplexen Argumente besteht. Durch die Substitution $y = ict$ hat man in (6,4) eine Verallgemeinerung der d'Alembertschen Wellengleichung: $u_{xx} - \frac{1}{c^2} u_{tt} = G(u) \cdot (u_x^2 - \frac{1}{c^2} u_t^2)$, und in (6,6), (6,7) deren allgemeine Lösung mit $\Phi = f_1(x - ct) + f_2(x + ct)$.

Versteht man unter Φ eine Partikularlösung der Potentialgleichung, so hat man in (6,7) eine Partikularlösung von (6,4).

Für unsere Gleichung (6,3) ist $G(u) = -\gamma$ und nach (6,6): $g(u) = \frac{1}{\gamma} e^{\gamma u}$.

Das gibt: $g^{-1}(\tau) = \frac{1}{\gamma} \ln(\gamma \tau)$. Also ist die allgemeine Lösung von (6,3):

$$\varphi = \frac{1}{\gamma} \ln(\gamma \Phi), \quad (6,8)$$

wenn Φ die allgemeine Lösung von $\Delta\Phi = 0$ ist. Die Feldstärke \mathfrak{E} wird:

$$\mathfrak{E} = -\frac{1}{\gamma} \text{grad } \ln(\gamma \Phi) = -\frac{1}{\gamma} \cdot \frac{\text{grad } \Phi}{\Phi}. \quad (6,9)$$

Daß hier Φ selbst auftritt, bedeutet nicht etwa eine Eichinvarianzschwierigkeit; Φ ist ja nicht das skalare Potential, sondern eine Lösung der Potentialgleichung. Unsere bisherigen Gleichungen sind alle eichinvariant, weil in ihnen nur die Feldstärken vorkommen, nicht die Potentiale selbst.

Wir interessieren uns zuerst für die kugelsymmetrische Lösung unseres Problems, nehmen also an, daß \mathfrak{E} radial gerichtet ist und nur von r abhängt. Dann hängt auch Φ nur von r ab und wir können für Φ die kugelsymmetrische Lösung

der Potentialgleichung einsetzen, also das Coulombsche Potential — bis auf eine additive Konstante, die man zu Φ hinzufügen kann. Es gilt ja bei Kugelsymmetrie:

$\Phi = \frac{b_1}{r} + b_2$. Man erhält so in Kugelkoordinaten r, ϑ, φ :

$$\mathfrak{E}_r = \frac{1}{\gamma r + \delta r^2}; \quad \delta = \frac{b_2 \gamma}{b_1}; \quad \mathfrak{E}_\vartheta = 0; \quad \mathfrak{E}_\varphi = 0, \quad (6,10)$$

was man unter der Voraussetzung der Kugelsymmetrie auch leicht aus $\text{div } \mathfrak{E} = \gamma \mathfrak{E}$ hätte finden können. δ ist eine noch zu bestimmende Konstante; man sieht, daß sie aus b_2 entstanden ist, und daß δ entscheidend ist für das Verhalten bei großen r . Allgemein sieht man an (6,9), daß die in Φ immer enthaltene additive Konstante das Verhalten in großer Entfernung bestimmt. In großer Entfernung von Singularitäten wird Φ konstant, und (6,9) zeigt also unmittelbar, daß die elektrische Feldstärke für große r durch $\text{grad } \Phi$ bestimmt ist, folglich mit der Feldstärke der klassischen Elektrodynamik übereinstimmt, weil Φ ja Lösung des unserem Fall entsprechenden klassischen Problems $\Delta \Phi = 0$ ist. Für große r kommt also beim Problem des ruhenden Teilchens mit radialsymmetrischem Feld das Coulombsche r^{-2} -Gesetz heraus, wie auch (6,10) lehrt. Es wird sich bald zeigen, daß auch die Konstante δ den vom Coulombschen Gesetz verlangten Wert bekommt.

Wir müssen ja mit der Lösung (6,10) die Endlichkeitsbedingungen (2,1), (2,2) erfüllen. Da unser Ansatz (6,2) für C zur Folge hat, daß:

$$\varrho = \gamma \mathfrak{E}^2, \quad (6,11)$$

so werden die beiden Bedingungen (2,1), (2,2) gleichbedeutend. Nennen wir die Ladung des Teilchens e , so wird:

$$\int \varrho d\tau_3 = e = \int \text{div } \mathfrak{E} d\tau_3 = 4\pi \int_0^\infty \frac{1}{r^2} \cdot \frac{d}{dr} (r^2 \mathfrak{E}_r) r^2 dr = 4\pi (r^2 \mathfrak{E}_r)_0^\infty = \frac{4\pi}{\delta}. \quad (6,12)$$

Andererseits ist wegen (6,1) und (6,11):

$$\int W d\tau_3 = m_0 c^2 = \frac{1}{2} \int \mathfrak{E}^2 d\tau_3 = \frac{1}{2\gamma} \int \varrho d\tau_3 = \frac{e}{2\gamma}. \quad (6,13)$$

Damit sind die Konstanten γ, δ bestimmt. Es ist:

$$\delta = \frac{4\pi}{e}; \quad \gamma = \frac{e}{2m_0 c^2} = \frac{r_0}{2e} \cdot 4\pi, \quad (6,14)$$

wenn wir wie in (1,1) setzen:

$$r_0 = \frac{e^2}{4\pi m_0 c^2}. \quad (6,15)$$

So wird aus (6,10):

$$\mathfrak{E}_r = \frac{e}{4\pi r^2 \left(1 + \frac{r_0}{2r}\right)}; \quad \mathfrak{E}_\vartheta = 0; \quad \mathfrak{E}_\varphi = 0, \quad (6,16)$$

als radialsymmetrische Lösung für ein ruhendes Teilchen, wenn wir den Ansatz (6,2) zugrunde legen. Man sieht, daß die aus Dimensionsgründen notwendige Konstante γ zwangsläufig zu fester Ladung e des Teilchens einen bestimmten Wert der Ruhenergie des Teilchens liefert, nämlich $\frac{e^2}{2\gamma}$, und einen scheinbaren Teilchenradius $r_0 = \frac{1}{2\pi} \cdot e\gamma$, der den Bereich bezeichnet, in dem die Feldstärke merklich von der Coulombschen verschieden ist. Das Teilchen ist keine Punkt-

singularität im Sinn des Coulombschen Feldes, hat vielmehr eine Ladungsverteilung ϱ , die aber im wesentlichen auf einen Bereich der Linearausdehnung r_0 beschränkt ist. Wir haben nämlich nach (6,11):

$$\varrho = \gamma \mathfrak{E}^2 = \frac{e}{8\pi r_0^2} \cdot \frac{\left(\frac{r_0}{r}\right)^4}{\left(1 + \frac{r_0}{2r}\right)^2}. \quad (6,17)$$

Die Raumladung des „Teilchens“ nimmt also für größere r ab wie r^{-4} . Für die innerhalb einer Kugel vom Radius a um $r = 0$ gelegene Ladung ergibt sich:

$$\varrho_a = \int_0^a \varrho r^2 dr \cdot 4\pi = \frac{e}{1 + \frac{r_0}{2a}}. \quad (6,18)$$

Innerhalb einer Kugel vom Radius r_0 liegt demnach $2/3$ der Teilchenladung e , für $a = 10 r_0$ wird $\varrho_a = 0,95 e$. Für ein Elektron wäre also innerhalb einer Kugel vom Radius $2,8 \cdot 10^{-13}$ cm um seinen Mittelpunkt 95% seiner Ladung e .

Durch (6,16) sind die Forderungen 1)–4) von § 2 erfüllt. Die Abweichung vom klassischen Verhalten beschränkt sich auf den Bereich $r \lesssim r_0$. Die Forderungen 6) und 7) sind schon durch den in § 4 entwickelten allgemeinen Ansatz erfüllt, auch 10), weil alle unsere Gleichungen lorentzinvariant sind; 10) brauchen wir deshalb nicht weiter zu erwähnen. Von 5) wird noch zu reden sein, 8) und 9) werden in dieser Arbeit nicht behandelt. Da die Theorie von vornherein relativistisch kovariant formuliert ist, macht das Auftreten der charakteristischen Länge r_0 keine Kovarianzschwierigkeiten, entgegen den Befürchtungen, die öfters ausgesprochen wurden angesichts des Problems, eine für das Teilchen charakteristische Länge in die Theorie aufzunehmen. Das Feld für ein gradlinig gleichförmig bewegtes Teilchen ergibt sich aus (6,16) durch Lorentztransformation.

In der Gleichung $\text{div } \mathfrak{E} = \gamma \mathfrak{E}^2$ stellt der Ausdruck $\gamma \mathfrak{E}^2$ die Korrektur gegenüber der klassischen Elektrodynamik dar. Es ist lehrreich, \mathfrak{E} mit der üblichen Methode der Störungsrechnung näherungsweise zu berechnen, also, salopp gesprochen, für \mathfrak{E} eine Entwicklung nach γ anzusetzen. An der strengen Lösung (6,16) sieht man aber, daß die Entwicklung gerade in dem Gebiet unbrauchbar werden muß, in dem man die Lösung vor allem haben möchte, nämlich für $r \lesssim r_0$. Denn gerade in diesem Gebiet bestimmt der Ausdruck $\gamma \mathfrak{E}^2$ das Verhalten der Feldstärke, und kann nicht mehr als kleine Größe behandelt werden. Man kann einen Ausdruck nicht als kleine Korrektur betrachten, der gerade da die anderen Einflüsse überwiegt, wo man das physikalische Verhalten kennenlernen will. Dementsprechend läßt sich aus der Näherungsrechnung nichts über das Verhalten von \mathfrak{E}_r für $r \ll r_0$ erfahren, und auch nichts drüber, ob die gesamte elektromagnetische Energie $\frac{1}{2} \int \mathfrak{E}^2 d\tau_3$ endlich bleibt oder nicht. Nach der Näherungsrechnung wäre das Integral divergent, die Gesamtenergie also unendlich, während die Rechnung mit dem strengen Ausdruck für \mathfrak{E} zeigt, daß die Gesamtenergie endlich ist, vergleiche (6,13). Wir deuten die Näherungsrechnung an; man setzt:

$$\mathfrak{E} = \mathfrak{E}^{(0)} + \mathfrak{E}^{(1)}; \quad \text{div } \mathfrak{E}^{(0)} = 0; \quad \text{div } \mathfrak{E}^{(1)} = \gamma(\mathfrak{E}^{(0)2} + \mathfrak{E}^{(1)2}) \approx \gamma \mathfrak{E}^{(0)2}; \quad (6,19)$$

und erhält:

$$\mathfrak{E}_r^{(0)} = \frac{e}{4\pi r^2}; \quad \mathfrak{E}_r^{(1)} = -\frac{\gamma e^2}{16\pi^2 r^3} = -\frac{r_0}{2r} \cdot \frac{e}{4\pi r^2}; \quad \mathfrak{E}_r \approx \frac{e}{4\pi r^2} \left(1 - \frac{r_0}{2r}\right), \quad (6,20)$$

was der Entwicklung von \mathfrak{E}_r in (6,16) nach $\frac{r_0}{2r}$ entspricht. Die Entwicklung konvergiert nur für $r > \frac{r_0}{2}$; das Integral $\int_0^\infty \mathfrak{E}_r^2 \cdot r^2 dr$ divergiert, mit der Näherungsfunktion (6,20) berechnet, und jedes neue Glied der Reihenentwicklung würde einen noch stärker divergierenden Ausdruck dazu liefern!

Die Kenntnis der allgemeinen Lösung (6,9) unseres elektrostatischen Problems:

$$\text{rot } \mathfrak{E} = 0; \quad \text{div } \mathfrak{E} = \gamma \mathfrak{E}^2; \quad \mathfrak{H} = 0 \quad (6,21)$$

erlaubt uns zu beweisen, daß es außer der Lösung (6,16) keine weitere gibt, welche endliche elektromagnetische Gesamtenergie und Ladung liefert, und nur einem ruhenden Teilchen entspricht, das in großer Entfernung Coulombsches Feld hat. Das Integral über die elektromagnetische Energie ist wegen (6,11) und nach dem Gaußschen Satz proportional zum Integral $\int \mathfrak{E}_n d\sigma$ über die Oberfläche des Bereiches. Da wir die Lösung untersuchen wollen, welche einem einzigen Teilchen entspricht, haben wir, wenn überhaupt, nur bei $r = 0$ eine Singularität zu erwarten. Das Integral $\int \mathfrak{E}_n d\sigma$ geht also über eine Kugel um $r = 0$ und über die unendlich ferne Kugel vom Radius R . Wir schreiben Φ in der allgemeinen Form an:

$$\begin{aligned} \Phi = & \sum_{l,m} (a_{lm} r^l + b_{lm} r^{l-1}) P_l^m(\cos \vartheta) \cos m \psi \\ & + \sum_{l,m} (c_{lm} r^l + d_{lm} r^{l-1}) P_l^m(\cos \vartheta) \sin m \psi. \end{aligned} \quad (6,22)$$

Da $r = 0$ der einzige singuläre Punkt von \mathfrak{E} ist, können wir annehmen, daß die Entwicklung von Φ im ganzen Bereich $0 < r < \infty$ konvergiert. Die Kugel um $r = 0$ liefert zum Integral $\int \mathfrak{E}_n d\sigma$ nichts, weil $(\mathfrak{E}_r \cdot r^2)_{r \rightarrow 0}$ nach (6,9) und (6,22) gleich null ist. Die Kugel mit dem Radius R liefert einen divergenten Beitrag für $R \rightarrow \infty$, wenn nicht alle a_{lm} und c_{lm} null sind bis auf a_{00} , das von null verschieden sein muß. Dann kann man \mathfrak{E} nach (6,9) für großes r in die Form setzen: $\mathfrak{E} \rightarrow -\frac{1}{\gamma a_{00}} \text{grad } \Phi$. Wenn dies mit dem Coulombschen Feld streng übereinstimmen soll, so muß $\text{grad } \Phi \sim \frac{1}{R^2}$ werden für großes r , und es müssen alle

b_{lm}, d_{lm} null sein bis auf b_{00} . Es bleibt wirklich nur $\Phi = a_{00} + \frac{b_{00}}{r}$, also die Lösung (6,16) übrig. Würde man die Forderung der Übereinstimmung mit dem Coulombschen Potential nicht stellen, so bestünde Φ aus Gliedern, welche ϑ und ψ enthalten, und würde im allgemeinen für Punkte zwischen $r = 0$ und $r = \infty$ null werden können, so daß dort \mathfrak{E} unendlich würde. Man wird auch so wieder auf (6,16) geführt.

Man kann den Beweis sogar auf den Fall erweitern, daß ein statisches Magnetfeld vorhanden ist; es braucht nur vorausgesetzt werden, daß überall $\mathfrak{v} = 0$, und daß das Feld $\mathfrak{E}, \mathfrak{H}$ zeitlich unveränderlich sei. Aus $\mathfrak{v} = 0$ folgt $\mathfrak{S} = 0$ wegen (5,4); das bedeutet $\mathfrak{H} = \lambda \cdot \mathfrak{E}$, wo λ eine skalare Funktion ist. Dann gelten die Gleichungen

$$\text{rot } \mathfrak{E} = 0; \quad \text{rot } \mathfrak{H} = 0; \quad \text{div } \mathfrak{H} = 0; \quad \text{div } \mathfrak{E} = \gamma (\mathfrak{E}^2 + \mathfrak{H}^2); \quad \mathfrak{H} = \lambda \cdot \mathfrak{E} \quad (66,23)$$

Das gibt:

$$\mathfrak{E} = -\text{grad } \varphi;$$

und:

$$\text{rot } \mathfrak{H} = -[\text{grad } \lambda, \text{grad } \varphi] = 0.$$

Also:

$$\text{grad } \lambda = \hbar \cdot \text{grad } \varphi,$$

wo \hbar ebenfalls eine skalare Funktion ist. Weiter:

$$\begin{aligned} -\text{div } \mathfrak{H} &= (\text{grad } \lambda, \text{grad } \varphi) + \lambda \cdot \Delta \varphi = 0; \quad -\text{div } \mathfrak{E} = \Delta \varphi = -\gamma (1 + \lambda^2) \cdot (\text{grad } \varphi)^2 \\ &= -\frac{1}{\lambda} (\text{grad } \lambda, \text{grad } \varphi) = -\frac{\hbar}{\lambda} (\text{grad } \varphi)^2. \end{aligned}$$

Es folgt:

$$\hbar = \gamma \lambda (1 + \lambda^2); \quad \text{grad } \lambda = \gamma \lambda (1 + \lambda^2) \text{grad } \varphi.$$

Das bedeutet:

$$\text{grad } \varphi = \frac{1}{2\gamma} \text{grad } \ln \left(\frac{\lambda^2}{1 + \lambda^2} \right); \quad \varphi = \frac{1}{2\gamma} \ln \frac{\lambda^2}{1 + \lambda^2}.$$

Die Konstante kann weggelassen werden. Aus der letzten Gleichung folgt:

$$1 + \lambda^2 = \frac{1}{1 - e^{2\gamma\varphi}}, \quad (6,24)$$

und durch Einsetzen in die Gleichung für $\Delta \varphi$:

$$\Delta \varphi = -\frac{\gamma}{1 - e^{2\gamma\varphi}} \text{grad}^2 \varphi. \quad (6,25)$$

Diese Gleichung ist vom Typ (6,4) und kann folglich allgemein integriert werden nach dem Schema von (6,6), (6,7). Die Funktion g von (6,6) wird hier:

$g(\varphi) = \frac{1}{\gamma k} \arcsin e^{\gamma\varphi}$; k ist eine Konstante. Die Lösung von (6,25) ist also:

$$\varphi = \frac{1}{\gamma} \ln \sin (\gamma k \Phi); \quad \Delta \Phi = 0. \quad (6,26)$$

Die Feldstärken werden:

$$\mathfrak{E} = -k \text{ctg} (k \gamma \Phi) \cdot \text{grad } \Phi; \quad \mathfrak{H} = \pm k \text{grad } \Phi; \quad \Delta \Phi = 0; \quad (6,27)$$

womit das Gleichungssystem (6,23) allgemein integriert ist. Für $k \rightarrow 0$ entsteht aus (6,27) wieder (6,9) und nur im Grenzfall $k \rightarrow 0$ ist die Lösung (6,27) brauchbar. Denn sonst würde \mathfrak{E} wegen des ctg -Faktors Unendlichkeitstellen im physikalischen Bereich $r > 0$ haben. (6,16) ist also die einzige statische Lösung für ein ruhendes Teilchen.

Es ist aber zu beachten, daß diese Schlüsse alle auf dem speziellen Ansatz (6,2) beruhen. Wir untersuchen jetzt den allgemeineren Ansatz:

$$iC = \delta (2A)^n, \quad n \geq 1. \quad (6,28)$$

Hier ist zu beachten, daß A nach Definition negativ ist, und daß iC reell sein muß. Es ist also genauer:

$$iC = \delta (-)^n \cdot \left(2 \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} \right)^n; \quad n \geq 1. \quad (6,29)$$

Wir schreiben:

$$(-)^n = e^{\pi i (2l+1)n}, \quad l \text{ ganz.}$$

Für δ schreiben wir: $\delta = \delta_1 + i\delta_2$, wo δ_1, δ_2 reelle Konstanten sind. Die Realitätsforderung für iC bedeutet dann:

$$\delta_1 \sin n\pi (2l+1) + \delta_2 \cos n\pi (2l+1) = 0. \quad (6,30)$$

Wir haben 2 Fälle zu unterscheiden:

a) $\sin n\pi(2l+1) = 0$; $\delta_2 = 0$; $n(2l+1) = m = \text{ganze Zahl}$. Es wird $\delta(-)^n = \delta_1(-)^m$. Man erhält im Fall a) also nur entweder δ_1 oder $-\delta_1$, je nach dem Wert von m , d. h. von n . Zum Fall a) gehören speziell alle ganzen Zahlen n .

Der Fall b) ist der allgemeine. Dann gilt:

b) $\sin n\pi(2l+1) \neq 0$; $n(2l+1) \neq \text{ganze Zahl}$; $\delta_1 = -\delta_2 \cdot \text{ctg } n\pi(2l+1)$.

Man erhält:

$$\delta(-)^n = -\frac{\delta_2}{\sin n\pi(2l+1)}. \quad (6,31)$$

Wenn der Ansatz (6,28) überhaupt physikalische Bedeutung hat, so muß er für alle Felder gelten, und n , δ sind für alle Felder gleich. Eine Festlegung von n , δ hätte man wohl erst von der quantentheoretischen Form dieser Theorie zu erwarten. Aber die Realitätsforderung gibt schon ohne Quantentheorie eine Einschränkung. Denn nach (6,30) hängt δ_1 mit δ_2 zusammen, und nach (6,30) hat $\delta(-)^n$ für festes δ_2 eine ganze Reihe von verschiedenen möglichen Werten, die sich zahlenmäßig und dem Vorzeichen nach unterscheiden, denn l ist eine beliebige ganze Zahl. Quantentheoretisch könnte demnach n gegeben sein; δ könnte eine Integrationskonstante sein, und δ_2 wäre dann durch eine physikalische Bedingung festzulegen. Die folgenden Betrachtungen zeigen für ein ruhendes Teilchen, daß $\delta(-)^n$ mit der Ruhmasse des Teilchens zusammenhängt. (6,31) bedeutet deshalb, daß wegen des noch willkürlichen l ganz verschiedene Werte der Ruhmasse des Teilchens möglich sind, und zwar im allgemeinen beliebig viele, diskrete Werte. Es gibt aber auch n -Werte, die nur eine endliche Anzahl von Massenwerten zur Folge haben: wenn n das Verhältnis einer ganzen Zahl zu einer geraden Zahl ist, (durch gemeinsame Faktoren ist gekürzt gedacht), dann beträgt die Anzahl der Massenwerte die Hälfte der geraden Zahl.

In diesem Zusammenhang ist erwähnenswert, daß wegen der Dimensionsgleichheit von A und \mathcal{E}^2 die Gln. (1,2) die Deutung zulassen ($E_0 = m_0 c^2$):

$$A^2 \sim E_0^2 C^6; \quad (6,32a)$$

$$A \sim r_0^2 C^2; \quad (6,32b)$$

$$A^2 \sim e^2 C^4. \quad (6,32c)$$

Wegen des Dimensionszusammenhangs $r_0 \sim \frac{e^2}{E_0}$ sind nur zwei von den drei Gln. (6,32) voneinander unabhängig. Durch Division der ersten zwei erhalten wir:

$$A^2 \sim \frac{E_0}{r_0} C^2, \quad (6,33)$$

oder auch: $A \sim \frac{E_0}{e} C$, was mit (6,2), (6,14) übereinstimmt, und mit (6,28) für $n = 1$.

Allgemein haben wir also die dimensionslosen Invarianten:

$$\frac{A}{C} \cdot \frac{e}{E_0}, \quad \frac{C^4 e^2}{A^2}, \quad \frac{D}{A}. \quad (6,34)$$

Die allgemeinste Form von (4,14) ist demnach:

$$A \left(\frac{A}{C} \cdot \frac{e}{E_0}, \quad \frac{C^4 e^2}{A^2}, \quad \frac{D}{A} \right) = 0, \quad (6,35)$$

wo A eine skalare Funktion bedeutet.

Für ein ruhendes Teilchen ist überall $\mathbf{v} = 0$, $\mathfrak{E} = 0$ und folglich:

$$-iC = \varrho = \frac{\delta_2 (2W)^n}{\sin n\pi (2l+1)}; \quad (6,36)$$

W ist immer > 0 , unter $(2W)^n$ ist der Hauptwert der Potenz zu verstehen.

$\frac{\delta_2}{\sin n\pi (2l+1)}$, das als Faktor bei W^n steht, hängt mit der Ruhenergie E_0 des Teilchens zusammen, wie wir noch zeigen werden. Das war ja schon im Falle $n = 1$ so, wie die Gleichungen (6,2) (6,13) lehren. Wenn der \sin Null wird und ein Teilchen endlicher Energie dargestellt werden soll, so ersetze man $\frac{\delta_2}{\sin n\pi (2l+1)}$

durch $-\frac{\delta_1}{\cos n\pi (2l+1)}$, siehe (6,30); damit ist der Fall a) in (6,36) mitumfaßt.

Daß der \sin verschiedenes Vorzeichen haben kann, bedeutet nach (6,36), daß durch den Ansatz (6,28) Teilchen entgegengesetzten Ladungsvorzeichens beschrieben werden. Die verschiedenen Werte des \sin geben die Möglichkeit, zu verstehen, daß es zu gegebener Gesamtladung Teilchen verschiedener Ruhmasse geben kann.

Der Ansatz (6,28) hat natürlich die in der allgemeinen Gl. (6,35) verlangte Form. Es ist:

$$\frac{\delta A^n}{C} = \left(\frac{A}{C} \cdot \frac{e}{E_0}\right)^{4n-3} \cdot \left(\frac{C^4 e^2}{A^3}\right)^{n-1} \cdot \text{const}; \quad \delta = \frac{r_0^{4n-3}}{e^{2n-1}} \cdot \text{const}. \quad (6,37)$$

(6,28) kann man schreiben:

$$-iC = \frac{\delta_2 e^{i\pi n (2l+1)}}{\sin n\pi (2l+1)} A^n. \quad (6,38)$$

Zur Erläuterung sei angenommen, daß $n = \frac{2k+1}{2}$ ist; k sei eine gegebene ganze Zahl. Das würde nach (6,28) (6,31) bedeuten:

$$-iC = \pm \delta_2 \cdot \left(2 \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}}\right)^{\frac{2k+1}{2}},$$

für ein ruhendes Teilchen deshalb:

$$-iC = \pm \delta_2 (2W)^{\frac{2k+1}{2}}.$$

Man hätte also in (6,28) die Beschreibung von Teilchen entgegengesetzten Ladungsvorzeichens. Die folgenden Überlegungen zeigen, daß es Teilchen gleichen Ladungsbetrages und gleicher Ruhmasse sind.

Wir rechnen für ein ruhendes Teilchen mit rein elektrischem Feld, lassen n beliebig (≥ 1) und führen die Abkürzung ein:

$$\frac{\delta_2}{\sin n\pi (2l+1)} \equiv \gamma. \quad (6,39)$$

Die Feldgleichungen geben:

$$\text{rot } \mathfrak{E} = 0; \quad \text{div } \mathfrak{E} = -iC = \gamma (\mathfrak{E}^2)^n; \quad n \geq 1. \quad (6,40)$$

Mit der Potenz rechts ist der Hauptwert gemeint. Wir untersuchen die radial-symmetrischen Lösungen. Die Ladung $\int \varrho d\tau_3$ des Teilchens nennen wir $\varepsilon \cdot e$; ε ist ≥ 0 definiert, $\varepsilon = \pm 1$. Da $(\mathfrak{E}^2)^n$ mit dem positiven Vorzeichen definiert ist, hat ϱ das Vorzeichen von γ , so daß $\varrho, \gamma, \varepsilon$ dasselbe Vorzeichen haben. Mit anderen Worten: ε ist das Vorzeichen von ϱ und von γ . Wir können also schreiben:

$$\varepsilon \gamma = |\gamma|. \quad (6,41)$$

Man bestätigt leicht, daß die radialsymmetrische Lösung von (6,40) lautet:

$$\mathfrak{E}_r = \frac{\varepsilon e}{4\pi r^2 \left(1 + \frac{|\gamma| (2n-1) (e/4\pi)^{2n-1}}{(4n-3)r^{4n-3}}\right)^{\frac{1}{2n-1}}} \quad (6,42)$$

Hier sind alle Potenzen als Hauptwerte gemeint; der Klammerausdruck im Nenner ist immer > 0 , es war $n \geq 1$ vorausgesetzt worden. Man sieht, daß für große r wieder das Coulombsche Gesetz herauskommt; in nächster Nähe von $r = 0$ dagegen ist $\mathfrak{E}_r \sim r^{-\frac{1}{2n-1}}$. Die „kritische Länge“ ist hier:

$$r_k = \left(\frac{|\gamma| (2n-1) \left(\frac{e}{4\pi}\right)^{2n-1}}{4n-3} \right)^{\frac{1}{4n-3}}, \quad (6,43)$$

wie man an (6,42) erkennt. Für $r > r_k$ geht das Feld um so schneller in das Coulombsche über, je größer n ist. Die Ladungsdichte ϱ wird:

$$\varrho = \gamma \cdot \frac{\left(\frac{e}{4\pi}\right)^{2n}}{r^{4n} \left(1 + \left(\frac{r_k}{r}\right)^{4n-3}\right)^{\frac{2n}{2n-1}}}; \quad (6,44)$$

sie fällt um so schneller ab mit wachsendem r , je größer n ist. Die Gesamtladung

ist endlich, und wegen $\text{div } \mathfrak{E} = \varrho$ gleich $4\pi \cdot \int_0^\infty \frac{1}{r^2} \frac{d}{dr} (r^2 \mathfrak{E}_r) \cdot r^2 dr = 4\pi (r^2 \mathfrak{E}_r)_0^\infty = \varepsilon e$, wie es sein muß. Die gesamte elektromagnetische Energie wird:

$$\int W d\tau_3 = 2\pi \int_0^\infty \mathfrak{E}_r^2 r^2 dr = \frac{e^2}{8\pi} \int_0^\infty \frac{dr}{r^2 \left(1 + \left(\frac{r_k}{r}\right)^{4n-3}\right)^{\frac{2n}{2n-1}}}. \quad (6,45)$$

Das Integral ist konvergent, der Integrand verhält sich an der unteren Grenze wie $r^{\frac{4(n-1)}{2n-1}}$, an der oberen wie r^{-2} . Das Integral läßt sich umformen; es wird, wenn man beachtet, das $\int W d\tau_3$ hier gleich der Ruhenergie E_0 des Teilchens ist:

$$E_0 = \frac{e^2}{8\pi r_k} \cdot J; \quad J = \int_0^\infty \frac{dx}{x^2 (1 + x^{-4n+3})^{\frac{2n}{2n-1}}}. \quad (6,46)$$

Das Integral J ist für $n \geq 1$ nur wenig von n abhängig; für $n = 1$ hat es den Wert 1, für $n \gg 1$ strebt es dem Wert $\frac{4}{3}$ zu. Man kann also schreiben:

$$E_0 \approx \frac{e^2}{8\pi r_k}. \quad (6,47)$$

Für das γ der Gl. (6,39) haben wir allgemein:

$$|\gamma| = \frac{(4n-3) \left(\frac{e}{4\pi}\right)^{1-2n}}{2n-1} \cdot r_k^{4n-3}. \quad (6,48)$$

Der Ansatz (6,28) war ein Versuch; der funktionale Zusammenhang zwischen C , A und D ist uns nicht bekannt. Sieht man (6,28) als gültig an, so kann man aus (6,39) schließen, daß es bei festem n , δ_2 verschiedene γ_i gibt. Die Ladung εe des Teilchens ist von γ_i unabhängig, nur das Vorzeichen der Ladung ist dasselbe.

wie das von γ_i . Zu festem Ladungsbetrag e gibt es nach (6,43) so viele r_k , und nach (6,46) so viele Ruhmassen m_0 , als es verschiedene $|\gamma_i|$ gibt. J hängt nur von n ab und ist daher für unsere Betrachtung eine Konstante. Zugleich sieht man, daß diese verschiedenen Ruhmassen zu verschiedenen Ladungsvorzeichen gehören können. Natürlich meine ich nicht, daß (6,28) die Lösung des Problems der verschiedenen Massen der Elementarteilchen sei. Aber es ist lehrreich zu sehen, daß mit diesem einfachen Ansatz die Existenz verschiedener Ruhmassen und Ladungsvorzeichen allein auf Grund einer Realitätsforderung herauskommt. Die Tatsache zu beschreiben, daß die Gesamtladung immer ein ganzes Vielfaches der Elementarladung ist, gehört wohl nicht in den Aufgabenbereich einer klassischen Theorie, ist vielmehr Aufgabe der Quantelung. Den Ladungsbetrag e betrachten wir als gegeben und untersuchen die Teilchen verschiedener Masse, die zum gleichen Ladungsbetrag gehören können. Wir diskutieren die Fälle $n = \frac{3}{2}$ und $n = \frac{6}{5}$. Für die Untersuchung des sin-Verhaltens in (6,39) genügt es, Werte für n zwischen 1 und 2 zu untersuchen; die um ganze Zahlen größeren n geben für $|\gamma_i|$ nichts Neues. Für die Zahlen n zwischen 2 und 3 würde das Vorzeichen von γ_i umgekehrt sein wie für die um 1 kleineren n ; das bedeutet Umkehrung der Ladungsvorzeichen.

$n = \frac{3}{2}$ gibt $\gamma_0 = -\delta_2$, $\gamma_1 = +\delta_2$; die übrigen l geben nichts Neues. Es gibt dann 2 Teilchenarten entgegengesetzt gleicher Ladung und gleicher Ruhmasse.

$n = \frac{6}{5}$ gibt für $l = 0$ bis 4 als sin-Argument: $\frac{\pi}{5} \cdot (6, 8, 0, 2, 4)$. $\frac{2\pi}{5}$ und $\frac{8\pi}{5}$ entsprechen 2 Teilchenarten gleicher Ruhmasse m_{01} , aber entgegengesetzter Ladung; $\frac{4\pi}{5}$, $\frac{6\pi}{5}$ gehören zu 2 anderen Teilchenarten mit gleicher Ruhmasse m_{02}

und entgegengesetzter Ladung. $0 \cdot \frac{\pi}{5}$ bedeutet wegen $\gamma \rightarrow \infty$, daß $E_0 \rightarrow 0$, $E_r \rightarrow 0$ geht, hat also keine physikalische Bedeutung. Für die Ruhmassen gilt:

$$\frac{m_{01}}{m_{02}} = \left| \frac{\sin \frac{2\pi}{5}}{\sin \frac{4\pi}{5}} \right|^{5/9} = 1,31:1.$$

Wenn n ein unechter Bruch aus größeren ganzen Zahlen oder eine irrationale Zahl wäre, so würden sich viele Ruhmassen ergeben, die sich erheblich unterscheiden könnten, weil der sin nahe an null liegen kann.

Im Fall $n = 1$ gibt es zu festem Ladungsbetrag nur eine Ruhmasse und nur ein Ladungsvorzeichen. Der Ansatz (6,2) ist also nicht geeignet, Teilchen verschiedener Ladung im gleichen Raumgebiet zu beschreiben. Er ist auch nicht geeignet, mehrere Teilchen der gleichen Art im gleichen Raumgebiet zu beschreiben. Wegen (6,9) gibt es nämlich z. B. für 2 Teilchen die statische Lösung, welche dem gewöhnlichen elektrostatischen Feld zweier Punktladungen entspricht:

$$\Phi = \frac{b_1}{r_1} + \frac{b_2}{r_2} + b_3; \quad \mathcal{E} = \frac{1}{\gamma} \cdot \frac{1}{\frac{b_1}{r_1} + \frac{b_2}{r_2} + b_3} \cdot \left(\frac{b_1 r_1}{r_1^2} + \frac{b_2 r_2}{r_2^2} \right); \quad (6,49)$$

formal steckt darin das Coulombsche Gesetz für große Entfernungen, wie wir ja schon wissen. Es kommt auch zahlenmäßig richtig heraus, wenn wir $\frac{4\pi b_1}{\gamma b_3}$ mit der Ladung e_1 des einen Teilchens gleichsetzen, $\frac{4\pi b_2}{\gamma b_3}$ mit der Ladung e_2 des anderen. Die Gesamtladung

$$\int \varrho d\tau_3 = \int \mathcal{E}_n d\sigma$$

wird gleich $e_1 + e_2$, wie es sein muß. Von dem Oberflächenintegral bleibt nur das über die unendlich ferne Kugel übrig; die Oberflächenintegrale um die Punkte $r_1 = 0$ und $r_2 = 0$ geben nichts. Die gesamte elektromagnetische Energie $\frac{1}{2} \int \mathfrak{E}^2 d\tau_3 = \frac{1}{2\gamma} \int \rho d\tau_3$ wird gleich $\frac{e_1 + e_2}{2\gamma}$. Nach (6,14) ist aber

$$\gamma = \frac{e_1}{2E_{01}} = \frac{e_2}{2E_{02}}.$$

So wird die gesamte elektromagnetische Energie einfach $E_{01} + E_{02}$, was bedeuten würde, daß keine potentielle Energie der Abstoßung zwischen den Teilchen gleichen Ladungsvorzeichens vorhanden ist. Dieses physikalisch unmögliche Ergebnis beruht 1. darauf, daß nach (6,2) $\int \mathfrak{E}_n d\sigma$ für ein elektrostatisches Feld proportional der Gesamtenergie ist, 2. darauf, daß zum Oberflächenintegral $\int \mathfrak{E}_n d\sigma$ nur die unendlich ferne Kugel etwas beiträgt, und 3. darauf, daß nach (6,9) die Feldstärke in großer Entfernung von den Teilchen sich geometrisch additiv aus den dort geltenden Coulombschen Feldstärken der 2 Teilchen zusammensetzt. Von diesen 3 Voraussetzungen wird die zweite allgemein gelten, sonst wäre die Forderung endlicher Gesamtenergie nicht erfüllbar. Die dritte ist ebenfalls allgemein gültig, die erste dagegen ist beschränkt auf den Fall $n = 1$. Für $n \neq 1$ ist $\int \mathfrak{E}_n d\sigma$, und damit die Gesamtladung nicht mehr proportional zur Gesamtenergie im statischen Feld, sondern proportional zu $\int (\mathfrak{E}^2)^n d\tau_3$. Die Gesamtladung bleibt gleich, unabhängig vom Abstand der Teilchen; sie ist nach unserer Theorie durch das Verhalten der Feldstärken in großer Entfernung von den Teilchen bestimmt, also durch das Coulombsche Gesetz.

Wir bemerken noch, daß wir (6,28) nur für den Fall rein elektrostatischen Feldes behandelt haben.

Eine für elektrostatische Felder mit (6,28) gleichwertige Gleichung würde sich für den Fall $n = \frac{3}{2}$ ergeben, wenn wir statt (6,28) allgemein ansetzen würden, daß die Invariante C^2 proportional sei zu $A^2 \cdot D$ oder zu $A \cdot D^2$. Wenn das magnetische Feld \mathfrak{H} null ist, werden beide Ausdrücke $\sim A^3$. Man hätte beim ersten Ansatz

$$iC = \pm \delta' (-A) \sqrt{\mathfrak{H}^2 - \mathfrak{E}^2}; \quad (6,50)$$

beim zweiten:

$$iC = \pm \delta'' \cdot \sqrt{-A \cdot (\mathfrak{H}^2 - \mathfrak{E}^2)}; \quad (6,51)$$

δ'' ist stets reell. Man sieht, daß die gebrochene Potenz $n = \frac{3}{2}$ durch einen einfachen Ansatz für C^2 herauskäme. Ob die Gleichungen (6,50), (6,51), welche für rein elektrostatische Felder Teilchen entgegengesetzt gleicher Ladung und gleicher Ruhmasse liefern, sonst physikalisch brauchbar sind, wäre zu untersuchen.

Aus den für ruhende Teilchen gewonnenen Ergebnissen lassen sich die für gradlinig gleichförmig bewegte Teilchen durch Lorentz-Transformation ableiten. Die Tatsache, daß für ein ruhendes Teilchen mit rein elektrischem Feld die beiden Invarianten A und D einander gleich werden:

$$D = \frac{1}{2} (\mathfrak{H}^2 - \mathfrak{E}^2) \rightarrow -\frac{1}{2} \mathfrak{E}^2; \quad A = -\frac{1}{2} \sqrt{W^2 - \frac{\mathfrak{E}^2}{c^2}} \rightarrow -\frac{1}{2} \mathfrak{E}^2,$$

läßt sich auch so aussprechen: Für ein gradlinig gleichförmig bewegtes Teilchen mit rein elektrischem Ruhfeld stimmen die Invarianten A und D überein.

§ 7. Der allgemeine Erhaltungssatz für Größen, die mit der Strömung \mathbf{v} wandern

Wir betrachten ein beliebig begrenztes Gebiet im Feld. Jedem Punkt dieses Gebietes entspricht ein bestimmter Wert von \mathbf{v} ; die Randpunkte des Gebietes wandern mit der ihnen zukommenden Geschwindigkeit. Dadurch verändert sich die Form und das Volumen des Gebietes. Die Volumenzunahme infolge der Strömung \mathbf{v} ist:

$$\frac{d}{dt} \int d\tau_3 = \int v_n d\sigma = \int \operatorname{div} \mathbf{v} d\tau_3. \quad (7,1)$$

Es sei nun B eine beliebige im Feld raumzeitlich veränderliche Größe. Wir untersuchen die Veränderung von $BV_4 d\tau_3$ infolge der Strömung:

$$\begin{aligned} \frac{d}{dt} (B V_4 d\tau_3) &= \left\{ \frac{\partial}{\partial t} (B V_4) + v_n (B V_4)_{,n} + B V_4 \operatorname{div} \mathbf{v} \right\} d\tau_3 \\ &= \left\{ \frac{\partial}{\partial t} (B V_4) + (B V_4 v_n)_{,n} \right\} d\tau_3 = i c (B V_4)_{,l} d\tau_3. \end{aligned} \quad (7,2)$$

B muß Tensoreigenschaft haben, wenn diese zeitliche Veränderung in kovarianter Weise einer anderen Größe gleich sein soll. Ist $(B V_4)_{,l} = 0$, so bleibt die in $d\tau_3$ enthaltene Menge von $B V_4$ trotz der Strömung erhalten. Ein Beispiel ist die Ladung; setzen wir B gleich der Invarianten C , so ist wegen $C V_l = S_l$ die rechte Seite von (7,2) null, und diese Gleichung spricht aus, daß die in $d\tau_3$ enthaltene Ladung $-i S_4 d\tau_3$ trotz der Strömung erhalten bleibt.

§ 8. Das Trägheitsgesetz und die Erhaltung der Ruhmasse

Wir ergänzen unseren Beweis für das Trägheitsgesetz und zeigen, daß sich bei fehlender äußerer Einwirkung die gradlinig gleichförmige Bewegung einstellt. In § 6 hatten wir bewiesen, daß sich die gradlinig gleichförmige Bewegung, wenn sie einmal vorhanden ist, ohne äußere Einwirkung nicht ändert. „Fehlende äußere Einwirkung“ definieren wir dadurch, daß das betrachtete Teilchen sein Feld unverzerrt mit sich führt. Wird in einem mitbewegten Koordinatensystem eine Änderung des Feldes feststellbar, so bedeutet dies Einwirkung einer äußeren Kraft.

Die Teilchen sind besondere Stellen im Feld; bei den Ansätzen für C , die wir in § 6 untersucht haben, waren die Teilchen Singularitäten des Feldes. Das mitbewegte System werden wir für ein Teilchen sinngemäß mit dieser besonderen Stelle des Feldes verbunden denken.

Wir beschränken uns auf Teilchenmodelle mit statischem Ruhfeld. Führt das Teilchen sein Feld unverzerrt mit, so ist im mitbewegten System überall $V_4 = 1$, $V_n = 0$. Das Beobachtersystem hat gegen dieses mitbewegte System die Geschwindigkeit \mathbf{v} ; vom Beobachtersystem aus gesehen werden alle Feldgrößen unverändert mit dem Teilchen mitgeführt. Es gilt also: $\frac{dA}{dt} = 0$, $\frac{dD}{dt} = 0$ und

wegen (4,14): $\frac{dC}{dt} = 0$. Andererseits ist $\frac{d\rho}{dt} = 0$, und deshalb wegen $\rho = -i S_4 = -i C V_4$ auch: $\frac{dV_4}{dt} = 0$. Aus (5,25) folgt dann auch: $\frac{dV_n}{dt} = 0$. Das bedeutet gradlinig gleichförmige Translation mit der gleichen Geschwindigkeit für alle Punkte des Feldes, weil im mitbewegten System überall $\mathbf{v} = 0$ ist. Unverzerrtes Mitführen eines statischen Feldes ist also gleichbedeutend mit gleichförmiger Translation desselben. Dieser Schluß beruht wesentlich darauf, daß die V_n durch T_{ki} und C , und damit durch die Augenblickswerte der Feldgrößen gegeben sind.

In der klassischen Theorie kann man nur schließen, daß ein gleichförmig gradlinig bewegtes Teilchen sein Feld unverzerrt mit sich führt, aber nicht, daß das Feld nur bei gleichförmiger Translation unverzerrt bleibt.

Dabei haben wir vorausgesetzt, daß das Ruhfeld statisch ist, und daß im Ruhesystem überall $v = 0$ ist. Wenn wir diese Voraussetzung beibehalten, können wir weiter schließen: aus $v = 0$ folgt $\mathfrak{S} = 0$. Für den Ansatz $C \sim A$ hatten wir in § 6 bewiesen, daß es nur eine einzige Lösung dieses Problems gibt. Nach der Einwirkung eines beliebigen äußeren Feldes haben die Teilchen also, sobald ihre Bewegung eine Translation geworden ist, wieder dieselbe Feldverteilung um sich wie vorher, ihre Ruhmasse ist daher erhalten geblieben. Wenn C nicht proportional zu A ist, gibt es im allgemeinen mehrere Lösungen für $v = 0$, und dementsprechend mehrere Teilchentypen; ein Beispiel ist $C \sim A^n$, $n > 1$. Nach der Feldeinwirkung ist einer dieser Typen realisiert. Das Teilchen ist durch die Einwirkung des äußeren Feldes entweder unverändert geblieben oder in einen anderen Teilchentyp übergegangen.

Eine andere Frage aber ist, ob dieser Zustand fehlender äußerer Einwirkung sich jemals wiederherstellt, wenn das Teilchen einmal durch eine äußere Kraft gestört worden ist. Denn unsere Definition kräftefreier Bewegung bedeutet unverzerrte Feldwanderung. Äußere Kraft ist in diesem Sinn auch das vom Teilchen bei Beschleunigung ausgesandte Feld; auch die klassische Theorie rechnete die Strahlungsrückwirkung zu den äußeren Kräften. Diese Frage gehört zum Problemkreis „Bewegungsgleichung“, zu dem wir in dieser Arbeit nur einige Bemerkungen machen wollen.

Daß ein gleichförmig gradlinig bewegtes Teilchen nicht strahlt, ist nach unserer Theorie wie nach der klassischen trivial. Wenn wir auf Ruhe transformieren, haben wir $v = 0$, $\mathfrak{S} = 0$, $\dot{W} = 0$; die vierte Komponente der Lorentzkraft G_4 ist auch Null, weil sie proportional ($v \mathfrak{E}$) ist. Allgemein folgt bekanntlich aus den Feldgleichungen (4,17), (4,18) der Erhaltungssatz:

$$T_{kn,n} + G_k = 0; G_k = S_i F_{ki}. \quad (8,1)$$

Seine vierte Komponente ($k = 4$) spricht die Energieerhaltung aus. Ein gleichförmig bewegtes Teilchen nimmt seine Energie mit, weil $\int \mathfrak{S}_n d\sigma = 0$ ist. Auch die Kraft des Teilchens auf sich selbst ist bei gleichförmiger Bewegung Null. Bildet man z. B. $\int \rho \mathfrak{E}_x d\tau_0$, so wird mit der Lösung (6,42), (6,44) das Integral proportional zu:

$$\int_{r=0}^{\infty} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \frac{dr \sin^2 \theta d\theta \cos \varphi d\varphi}{r^{4n} \left(1 + \left(\frac{r_k}{r}\right)^{4n-3}\right)^{\frac{2n+1}{2n-1}}};$$

es ist Null wegen der Integration über φ ; das Integral über r allein ist konvergent.

Obwohl die räumlichen Komponenten G_i der Lorentzkraft an den einzelnen Raumstellen im allgemeinen nicht Null sind, entsteht keine Ladungsbewegung, v ist ja überall Null. Das Teilchen hält zusammen und ändert sein Feld nicht, obwohl $G_i \neq 0$ ist.

Grundsätzlich sind auch Teilchenmodelle möglich, bei denen im Ruhesystem nicht überall $v = 0$ ist. Wie das Trägheitsgesetz und die Erhaltung der Ruhmasse dann lautet, wäre noch zu untersuchen. Auch die Identifizierung von $-A V_4$ mit

der Ruhenergiedichte ist eine Annahme, die geändert werden könnte; sie hängt nicht zwangsläufig mit unserer Verknüpfung von S_n und T_{kn} zusammen, und auch nicht mit den Feldgleichungen, wie wir in § 4 betonten.

§ 9. Allgemeine Erhaltungssätze

Wenn K_{nl} ein antisymmetrischer Tensor ist:

$$K_{nl} = -K_{ln}, \quad (9,1)$$

und B eine beliebige Tensorgröße, so gilt in allen Bezugssystemen die Identität:

$$(K_{nl} B)_{,n,l} \equiv 0, \quad (9,2)$$

denn beim Vertauschen von n mit l ändert der Ausdruck links sein Vorzeichen. (9,2) hört auf, eine Identität zu sein und wird zu einer Gleichung, sobald in (9,2) Beziehungen eingesetzt werden, denen K_{nl} oder B genügen.

Setzt man z. B. K_{nl} gleich dem Feldstärkentensor F_{nl} , und benützt die Feldgleichung (4,17), so wird aus (9,2) der Erhaltungssatz:

$$(F_{nl} B_{,n} - S_l B)_{,l} = 0. \quad (9,3)$$

Da B ein beliebiger Tensor sein kann, können wir B gleich dem Viererpotential Φ_k setzen, wenn wir wollen. Dann entsteht aus (9,3) der Erhaltungssatz (8,1) für den Energie-Spannungstensor, wie wir jetzt zeigen wollen. Wegen $S_{,l} = 0$ und (4,20) folgt aus (9,3):

$$S_l \Phi_{k,l} = (F_{nl} \Phi_{k,n})_{,l} = -S_l F_{kl} + S_l \Phi_{l,k} = -G_k + S_l \Phi_{l,k} = (F_{nl} F_{nk} + F_{nl} \Phi_{n,k})_{,l};$$

und weiter wegen (4,17):

$$G_k + (F_{nl} F_{nk})_{,l} = S_l \Phi_{l,k} - S_n \Phi_{n,k} - F_{nl} \Phi_{n,k,l} = -1/2 F_{nl} \Phi_{n,l,k} - 1/2 F_{ln} \Phi_{l,n,k} \\ = -1/2 F_{nl} F_{ln,k} = 1/4 (F_{nl} F_{nl})_{,k} = 1/4 \delta_{kl} (F_{nj} F_{nj})_{,l},$$

was zu beweisen war.

(9,2) beruht auf der Antisymmetrie von K_{nl} . Es gibt neben (9,2) noch andere Identitäten, bei denen gar nichts über die darin vorkommenden Tensoren vorausgesetzt werden muß. Wir bilden mit beliebigen Vierervektoren A_n, B_l die Größen:

$$R_{kl} \equiv [(A_n B_l)_{,k} - \delta_{kl} (A_n B_n)_{,j}]_{,n}; \quad (9,4)$$

$$H_{kl} \equiv [(A_n B_k)_{,l} - (A_l B_k)_{,n}]_{,n} \quad (9,5)$$

und die durch Vertauschen von A mit B aus (9,4) und (9,5) entstehenden Größen:

$$X_{kl} \equiv [(A_l B_n)_{,k} - \delta_{kl} (A_l B_n)_{,j}]_{,n}; \quad I_{kl} \equiv [(A_k B_n)_{,l} - (A_k B_l)_{,n}]_{,n}. \quad (9,6)$$

Es gilt identisch:

$$R_{kl,l} \equiv 0; \quad H_{kl,l} \equiv 0; \quad X_{kl,l} \equiv 0; \quad I_{kl,l} \equiv 0 \quad (9,7)$$

für beliebige Vierervektoren A_n, B_l . Zum Beweis genügt das Anschreiben der nach (9,7) zu bildenden Differentiationen. (9,7) bleibt richtig, wenn man $A_n \cdot B_l$ ersetzt durch einen beliebigen Tensor L_{nl} zweiter Stufe.

Die Tensoren $R_{kl}, H_{kl}, X_{kl}, I_{kl}$ haben im allgemeinen keine Symmetrieeigenschaften. Einen symmetrischen Tensor Y_{kl} erhält man durch den Ansatz:

$$Y_{kl} \equiv R_{kl} + X_{kl} + H_{kl} + I_{kl}; \quad Y_{kl} \equiv Y_{lk}; \quad Y_{kl,l} \equiv 0. \quad (9,8)$$

Y_{kl} ist der einzige symmetrische Tensor, der aus (9,4)–(9,6) gebildet werden kann. Auch in (9,8) ist der Erhaltungssatz (8,1) enthalten. Man setze $A_n \equiv B_n \equiv \Phi_n$; dann wird $R_{kl} \equiv X_{kl}, H_{kl} \equiv I_{kl}$. So liefert (9,8):

$$Z_{kl} \equiv [(\Phi_n \Phi_l)_{,k} + (\Phi_n \Phi_k)_{,l} - (\Phi_k \Phi_l)_{,n} - \delta_{kl} (\Phi_n \Phi_n)_{,n}]_{,n}; \quad Z_{kl} \equiv Z_{lk}; \quad Z_{kl,l} \equiv 0. \quad (9,9)$$

Die Umformung von Z_{kl} wird am einfachsten, wenn wir (4,21) benützen. Dann wird:

$$\left. \begin{aligned} Z_{kl} &= (\Phi_n \Phi_{l,n})_{,k} + (\Phi_n \Phi_{k,n})_{,l} - (\Phi_{k,n} \Phi_l + \Phi_k \Phi_{l,n})_{,n} \\ &\quad - \delta_{kl} \Phi_{n,j} \Phi_{j,n} = \Phi_n (\Phi_{l,n,k} + \Phi_{k,n,l}) + (\Phi_{n,k} \Phi_l)_{,n} + (\Phi_{n,l} \Phi_k)_{,n} \\ &\quad - (\Phi_{k,n} \Phi_l + \Phi_k \Phi_{l,n})_{,n} - \delta_{kl} \Phi_{n,j} \Phi_{j,n} = (F_{kn} \Phi_l + F_{ln} \Phi_k)_{,n} \\ &\quad + \Phi_n (\Phi_{l,n,k} + \Phi_{k,n,l}) - \delta_{kl} \Phi_{n,j} \Phi_{j,n} = S_k \Phi_l + S_l \Phi_k + 2 F_{kn} F_{nl} \\ &\quad + F_{kn} \Phi_{n,l} + F_{ln} \Phi_{n,k} + \Phi_n (\Phi_{l,n,k} + \Phi_{k,n,l}) - \delta_{kl} \Phi_{n,j} \Phi_{j,n}. \end{aligned} \right\} (9,10)$$

Aus $Z_{kl,l} = 0$ folgt so:

$$\left. \begin{aligned} Z_{kl,l} &= 0 = S_{k,l} \Phi_l + S_l \Phi_{k,l} - 2 (F_{nk} F_{nl})_{,l} + F_{kn,l} \Phi_{n,l} - F_{kn} S_n \\ &\quad - S_n \Phi_{n,k} + F_{ln} \Phi_{n,k,l} + \Phi_{n,l} (\Phi_{l,n,k} + \Phi_{k,n,l}) - \Phi_n S_{k,n} - 2 \Phi_{n,j} \Phi_{j,n,l} \\ &\quad = -2 (G_k + T_{kl,l}) - \frac{1}{2} \cdot 2 \cdot F_{nj} F_{nj,k} + 2 \Phi_{n,l} \Phi_{n,k,l} \\ &\quad - 2 \Phi_{n,j} \Phi_{j,n,k} = -2 (G_k + T_{kl,l}), \end{aligned} \right\} (9,11)$$

was zu beweisen war.

Aus (9,4)–(9,6) kann man auf eine einzige Weise auch einen antisymmetrischen Tensor U_{kl} bilden, dessen Divergenz Null ist:

$$U_{kl} \equiv R_{kl} - X_{kl} - H_{kl} + I_{kl}; \quad U_{kl} \equiv -U_{lk}; \quad U_{kl,l} \equiv 0. \quad (9,12)$$

Die Sätze (7,1), (7,2) können wir jetzt allgemein fassen. Wir betrachten eine beliebige Tensorgröße K – in (7,2) hieß sie $B V_4$ – und fragen nach der zeitlichen Veränderung von $K d\tau_3$, wenn die Grenzen von $d\tau_3$ mit der Geschwindigkeit w wandern. Wir bilden:

$$w_l \equiv \frac{1}{ic} w_l; \quad l \neq 4; \quad (9,13)$$

dann gilt:

$$\frac{d}{dt} (K d\tau_3) = ic (K_{,4} + w_l K_{,l} + K w_{l,l}) d\tau_3 = ic (K_{,4} + (K w_l)_{,l}) d\tau_3. \quad (9,14)$$

Wir definieren:

$$W_i \equiv \{w_i, l\} \cdot W_4; \quad (9,15a)$$

das gibt:

$$W_l W_l = \left(1 - \frac{w^2}{c^2}\right) W_4^2. \quad (9,15b)$$

Wenn W_l eine Vierergeschwindigkeit $\frac{dx_l}{d\tau} = \frac{dx_l}{ic dt \sqrt{1 - \frac{w^2}{c^2}}}$ sein soll, so folgt:

$$W_4 = \frac{1}{\sqrt{1 - \frac{w^2}{c^2}}}; \quad W_l W_l = 1. \quad (9,16)$$

Aus (9,14) wird:

$$\frac{d}{dt} (K d\tau_3) = ic \left(\frac{K W_l}{W_4} \right)_{,l} \cdot d\tau_3. \quad (9,17)$$

Man sieht daraus, daß zu jedem Erhaltungssatz eine Strömungsgeschwindigkeit W_l gehört.

Zum Beispiel entspricht in: $Y_{kl,l} = 0$ die Größe Y_{k4} unserem K_l und deshalb ist:

$$Y_{kl} = K \frac{W_l}{W_4} = Y_{k4} \frac{W_l}{W_4}.$$

Da die Y verschiedenen Index k haben können, wird man genauer schreiben:

$$W_i^{(k)} = \frac{Y_{ki}}{Y_{k4}} \cdot W_4^{(k)}. \quad (9,18)$$

Aus (9,16) findet man:

$$W_4^{(k)} = \sqrt{\frac{Y_{k4}^2}{Y_{(k)1} Y_{(k)1}}}; \quad (9,19)$$

damit ist $W_i^{(k)}$ völlig bestimmt. Wandern die Grenzen von $d\tau_3$ mit der Geschwindigkeit:

$$w_i^{(k)} = \frac{ic Y_{ki}}{Y_{k4}}, \quad (9,20)$$

so bleibt $\int Y_{k4} d\tau_3$ zeitlich konstant. Verschwindet $w_i^{(k)}$ am Rand des Gebietes, dann wandern die Grenzen des Gebietes nicht, und das Integral bleibt in einem räumlich unveränderlichen Gebiet zeitlich konstant.

Dieselbe Betrachtung wie für Y_{ki} läßt sich für jede Tensorgröße machen. Nimmt man die Gleichung $S_{i,l} = 0$ als Beispiel, so liefert (9,20):

$$w_l = \frac{ic S_{li}}{S_4} = v_l,$$

und (9,19):

$$W_4 = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}},$$

wie es sein muß.

Im ladungsfreien Raum ist $T_{ki,l} = 0$; aus (9,20) wird: $w_l^{(k)} = \frac{ic T_{kl}}{T_{k4}}$. Speziell ist: $w_l^{(4)} = \frac{\mathcal{E}_l}{W}$, worin auch das bekannte Ergebnis: $cW = |\mathcal{S}|$ für fortschreitende ebene Wellen enthalten ist.

§ 10. Bemerkungen zur wellenmechanischen Deutung

In dieser Arbeit soll nicht versucht werden, die hier gegebenen Ansätze einer Elektrodynamik systematisch in die Quantentheorie umzuschreiben; es soll nur eine naheliegende quantentheoretische Deutung für das Modell des ruhenden Teilchens (§ 6) gegeben werden. Wir beschränken uns dabei auf den Fall $n = 1$; die Rechnungen werden dann einfacher, qualitativ gelten die so gefundenen Ergebnisse auch für $n > 1$.

Weil nur ein Teilchen vorhanden ist, gibt $\int \rho d\tau_3 = e$ die Ladung dieses Teilchens; $\frac{\rho}{e}$ kann in diesem Fall als Aufenthaltswahrscheinlichkeit gedeutet werden.

Da $\frac{\rho}{e}$ auf die nächste Nähe ($r \approx r_0$) von $r = 0$ beschränkt ist, hat man ein Teilchen, von dem man mit sehr großer Genauigkeit weiß, daß es sich bei $r = 0$ befindet. In der gewöhnlichen Wellenmechanik hätte man zur Beschreibung dieses Teilchens ein Wellenpaket zu verwenden, das praktisch eine Zackenfunktion im Lagenraum darstellt; wellenmechanisch ist ja r_0 eine vernachlässigbare Länge. In der Wellenmechanik hat man dann praktisch gleiche Wahrscheinlichkeit für alle Impulse. Unsere Lösung muß qualitativ dieselbe Eigenschaft haben für den Bereich, in dem die Wellenmechanik gilt, also für $\lambda \gg r_0$.

Wir betrachten $\sqrt{\varrho}/e$ als Wahrscheinlichkeitsamplitude:

$$\sqrt{\frac{\varrho}{e}} = \sqrt{\frac{r_0}{8\pi}} \cdot \frac{1}{r(r + \frac{r_0}{2})}, \quad (10,1)$$

und zerlegen sie in de Broglie-Wellen:

$$\sqrt{\frac{r_0}{8\pi}} \cdot \frac{1}{r(r + \frac{r_0}{2})} = \frac{1}{(2\pi)^{3/2}} \int a(\mathbf{k}) e^{i(\mathbf{k}\mathbf{r})} d\mathbf{k}. \quad (10,2)$$

$|a|^2 d\mathbf{k}$ ist die Wahrscheinlichkeit, den Impuls $\hbar \mathbf{k}$ im Bereich $d\mathbf{k}$ durch Messung zu finden, bei einem Teilchen, das sich praktisch an der Stelle $r = 0$ befindet. Es wird:

$$\begin{aligned} a(\mathbf{k}) &= \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{r_0}{8\pi}} \int \frac{e^{-i(\mathbf{k}\mathbf{r})}}{r(r + \frac{r_0}{2})} d\mathbf{r} = \frac{\sqrt{r_0}}{8\pi^2} \int \frac{e^{-i\mathbf{k}\mathbf{r}\cos\theta}}{r + \frac{r_0}{2}} \cdot r dr \sin\theta d\theta d\varphi \\ &= \frac{\sqrt{r_0}}{2\pi k} \int_0^\infty \frac{\sin x}{x + x_0} dx; \quad x_0 = \frac{k r_0}{2} = \frac{r_0}{2\lambda}; \quad \lambda = \frac{2\pi}{k} = \frac{1}{k}. \end{aligned} \quad (10,3)$$

Das Integral:

$$J = \int_0^\infty \frac{\sin x}{x + x_0} dx \quad (10,4)$$

läßt sich auf den Integralsinus und -cosinus zurückführen:

$$J = \left(\frac{\pi}{2} - \text{Si } x_0\right) \cdot \cos x_0 + \text{Ci } x_0 \cdot \sin x_0. \quad (10,5)$$

Der gewöhnlichen Wellenmechanik entspricht der Fall $x_0 \ll 1$. Dann ist:

$$J \approx \frac{\pi}{2} + x_0 \ln x_0 \approx \frac{\pi}{2}; \quad a(\mathbf{k}) \approx \frac{\sqrt{r_0}}{4k}. \quad (10,6)$$

Die Wahrscheinlichkeit, einen festen Impulsbetrag $\hbar k$ zu finden, ist:

$$k^2 dk \cdot \int |a|^2 \sin\theta d\theta d\varphi = \frac{r_0 \pi}{4} dk; \quad (10,7)$$

sie ist von k unabhängig. Dies ist das Gebiet nicht zu großer Impulse: $\hbar k \ll \frac{2\hbar}{r_0}$.

Interessanter ist das Gebiet sehr hoher Impulse: $x_0 \gg 1$. Dort ist $\hbar k \gg \frac{2\hbar}{r_0}$. Dann gilt:

$$J \approx \frac{1}{x_0}; \quad a(\mathbf{k}) \approx \frac{1}{\pi k^2 \sqrt{r_0}}. \quad (10,8)$$

Die Wahrscheinlichkeit, einen festen Impulsbetrag $\hbar k$ zu finden, wird:

$$\frac{4}{\pi k^3 r_0} dk, \quad (10,9)$$

also um so kleiner, je größer der Impuls ist. Sehr hohe Impulse kommen im Impulsspektrum nicht merklich vor.

(0,1)

(0,2)

Mes-
det.

(0,3)

(0,4)

(0,5)

ist:

(0,6)

(0,7)

$\frac{\hbar}{\epsilon_0}$
 $\frac{\hbar}{\epsilon_0}$

(0,8)

(0,9)

m

Die

Ge
sch
me
we

ist
ll:

Die
—
die
v e
gon
lös
po

da
Wi
ent
an
gil

wi
in

Bei $x_0 = \pi$ ist $\hbar k = \frac{\hbar}{r_0}$; $\lambda = r_0$, und:

$$J(\pi) = 0,28; a(t) = \frac{r_0^{3/2} \cdot 0,28}{4\pi^2}. \quad (10,10)$$

Die Wahrscheinlichkeit, diesen Impuls wert $\frac{\hbar}{r_0}$ zu finden, ist:

$$\frac{0,28^2 \cdot r_0}{\pi} dk. \quad (10,11)$$

Gegenüber dem Fall der gewöhnlichen Wellenmechanik ($x_0 \ll 1$) ist diese Wahrscheinlichkeit um den Faktor $\frac{0,28^2 \cdot 4}{\pi^2} \approx 0,03$ kleiner. Der Abfall vom wellenmechanischen Verhalten der Wahrscheinlichkeit zu den sehr kleinen Werten, welche den hohen Impulsen zugehören, geschieht im Gebiet $x_0 \approx 1, \lambda \approx r_0, \hbar k \approx \frac{\hbar}{r_0}$.

Zusatz bei der Korrektur: Zu § 3 und 5: Die Matrix \mathfrak{T} des Tensors T_{kn} ist reell, symmetrisch und hermitisch. Wenn $A \neq 0$, kann man die Matrix $\mathfrak{U}: U_{kn} = \frac{T_{kn}}{\sqrt{A}}$ einführen; sie ist wegen Satz IX von § 3 unitär und orthogonal.

Die Matrix \mathfrak{T} hat nach Satz X von § 3 die 2 doppelten Eigenwerte $+A$ und $-A$, die Matrix \mathfrak{U} also die 2 doppelten Eigenwerte ± 1 . Die Größen V_i sind die „Eigenlösungen“ von \mathfrak{U} , sie sind durch die Gleichung (3,6) normiert. Die zu verschiedenem Eigenwert gehörigen Eigenlösungen V_i sind zueinander orthogonal. Die zum Eigenwert $+A$ von \mathfrak{T} , bzw. $+1$ von \mathfrak{U} , gehörigen Eigenlösungen $V_i^{(1)}, V_i^{(2)}$ sind dann zueinander orthogonal, wenn für ihre vierte Komponente V_4 gilt:

$$V_i^{(1)2} + V_i^{(2)2} = \frac{T_{44} + A}{2A};$$

dabei ist [vgl. (5,14)]: $\Delta_1 \neq 0, A \neq 0$ vorausgesetzt, außerdem müssen die Wurzelvorzeichen in (5,16) so gewählt werden, daß das Produkt dieser Vorzeichen entgegengesetzt gleich dem Produkt der Vorzeichen von $V_4^{(1)} \cdot V_4^{(2)}$ ist. Da die anderen 2 Eigenlösungen zum Eigenwert $-A$ von \mathfrak{T} bzw. -1 von \mathfrak{U} gehören, gilt für sie:

$$V_i^{(3)2} + V_i^{(4)2} = -\frac{T_{44} - A}{2A};$$

wieder ist $\Delta_1 \neq 0, A \neq 0$ vorausgesetzt, und das Produkt der Wurzelvorzeichen in (5,16) muß entgegengesetzt gleich dem Produkt der Vorzeichen von $V_4^{(3)} \cdot V_4^{(4)}$ sein.

Mainz, Institut für Theoretische Physik der Universität.

(Bei der Redaktion eingegangen am 20. April 1950.)

Ausleuchtung von Phosphoren durch Ultra-Schall

Von A. Eckardt und O. Lindig

(Mit 4 Abbildungen)

Inhaltsübersicht

Die Ausleuchtung von langsam abklingenden Phosphoren durch Einstrahlung von Ultraschall wird benutzt, um mittels Schirmbildphotographie das Schallfeld eines Strahlungsquarzes sichtbar zu machen. Die zu diesem Verfahren notwendigen Intensitäten übertreffen größenordnungsmäßig die zur Bilderzeugung mittels des von Pohlmann angegebenen Schwebeteilchenverfahrens erforderlichen.

Die Erzeugung einer ganz schwachen Lumineszenz durch Ultraschall in gasartigem Wasser wurde bereits 1935 von H. Frenzel und H. Schultes¹⁾ beobachtet. Ebenso sind Versuche bekannt²⁾, durch Beimischen von Phosphoren in die schalleitende Flüssigkeit, den Schallstrahl sichtbar zu machen. Hier wurde nun versucht, auf einem Leuchtschirm, wie er auch in der Röntgentechnik üblich ist, das von einem Schwingquarz erzeugte Schallfeld sichtbar zu machen. Phosphore, deren Leuchten langsam abklingt, kann man bekanntlich durch Zuführen von Energie, etwa durch Infrarot-Bestrahlung, zur schnellen Abgabe der aufgespeicherten Quantenenergie zwingen. Dasselbe muß auch möglich sein, wenn man eine entsprechende Energiezuführung durch Ultraschall erreicht. Es wurden daher Phosphore verwendet, welche nach Erregung durch UV langsam abklingen. Sehr gut geeignet sind entsprechend aktivierte Zn—Cd—S-Phosphore. Sobald auf einen derartigen, auf einem Schirm angebrachten Phosphor, Ultraschall fiel, leuchteten die getroffenen Stellen hell auf, so daß sich das Bild des Ultraschallfeldes scharf vom Abklingleuchten der Umgebung abhob. Nach dem Abschalten des Ultraschalles war der vorher hell leuchtende, von diesem getroffene Teil des Leuchtschirmes als eine gegenüber dem Abklingleuchten dunklere Stelle sichtbar, man erhielt also dann für das vorher positive Bild ein negatives.

Die Versuchsanordnung bestand aus einem mit vollkommen luftfreien Wasser gefüllten Versuchstrog, dessen Stirnseite eine Abschlußfolie aus Zelluloid trug, welcher gegenüber in verstellbarer Entfernung von 12–6 cm ein schallabstrahlender Straubelquarz angebracht war. Die Außenseite dieser Folie wurde mit dem zu untersuchenden Phosphor bedeckt, stellte also den Leuchtschirm dar. Die Eingangsleistung in den Schallgeber betrug etwa 70 Watt, die Frequenz des abgestrahlten Schalles 1,42 MHz. Der Leuchtschirm wurde 1 Minute über einen Quarzkondensor mit einer Bogenlampe beleuchtet, etwa $\frac{1}{2}$ Minute nach Abschalten der

¹⁾ H. Frenzel u. H. Schultes, Z. physik. Chem. B **27**, 421 (1934).

²⁾ L. Chamber, Physic. Rev. **49**, 881 (1936).

Bogenlampe wurde der Schallgeber erregt und das Bild des Schallfeldes photographiert. Es gelang mit Belichtungszeiten von 6–12 Sekunden bei einer Lichtstärke des Objektives von $f:4,0$ auf Agfa-Fluorapidfilm gut durchbelichtete Aufnahmen zu bekommen. Nach dem Abschalten des Quarzes wurde das negative Bild des Schallfeldes, d. h. der ausgeleuchtete Teil des Schirmes, ebenfalls aufge-

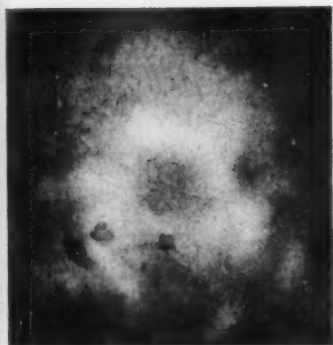


Abb. 1a. Abstand: Quarz-Schirm etwa 10 cm 30 s nach Ende der UV-Bestrahlung. Während der Beschallung 16 s belichtet

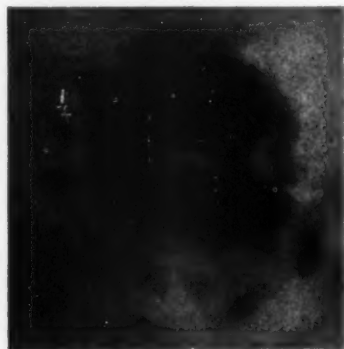


Abb. 2a. Wie Abb. 1a, aber nach Abschalten der Schallquelle. Belichtungszeit 35 s



Abb. 1b. Abstand: Quarz-Schirm etwa 8 cm, sonst wie Abb. 1a

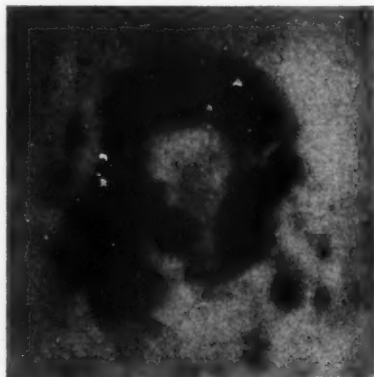


Abb. 2b. Wie Abb. 1b nach Abschalten der Schallquelle. Belichtungszeit 40 s

nommen, allerdings waren hierzu etwas längere Belichtungszeiten von 20–30 Sekunden notwendig. In den Abb. 1 und 2 sind entsprechende Aufnahmen gezeigt.

Die merkwürdige wolkige Struktur des Schallfeldbildes hat verschiedene Ursachen: Einerseits ist die Abstrahlung vom Quarz durchaus nicht gleichmäßig, andererseits machen sich Interferenzerscheinungen und stehende Wellen bemerkbar,

was man besonders daran erkennt, daß durch geringes Wölben der Folie, wenn man etwa mit dem Finger auf den Schirm drückt, die dunklen Stellen verschoben werden können. Auch ist zu berücksichtigen, daß der Schirm mit Laboratoriumsmitteln hergestellt war und keineswegs eine gleichmäßige Schichtdicke des Fluoreszenzstoffes aufwies.

Es soll der Zweck laufender Untersuchungen sein, den Mechanismus der Phosphorausleuchtung durch Ultraschall aufzuklären. Insbesondere soll der zeitliche Verlauf der Erscheinung untersucht werden. Betreffend die Anwendung des Schirmbildverfahrens zur Sichtbarmachung biologischer und technischer Objekte ist zunächst zu sagen, daß, soweit es die hier untersuchten Phosphore anbetrifft, die Empfindlichkeit des Verfahrens weit zurücksteht gegenüber dem von Pohlmann³⁾ angegebenen Schallbildwandler. Die hier zur Sichtbarmachung verwendeten Schallintensitäten dürften in vielen Fällen bereits biologisch schädigend wirken, bei Dazwischenschaltung dicker zu untersuchender Materieschichten dagegen dürfte die auf den Schirm auffallende Intensität zum Nachweis zu schwach sein.

³⁾ R. Pohlmann, Z. angew. Physik 1, 181 (1948).

Anmerkung während der Korrektur: Inzwischen erhielten wir eine Mitteilung von H. Schreiber und W. Degner vom Institut für Strahlungsforschung, Berlin, daß es dort in gleicher Weise gelungen ist, kontaktphotographische Bilder der Strahlungsfelder von Rundquarzen zu erhalten. Wir möchten auch an dieser Stelle für die freundliche Überlassung des Manuskriptes danken.

Jena, Technisch-Physikalisches Institut der Friedrich-Schiller-Universität.

(Bei der Redaktion eingegangen am 17. April 1950.)

Prozesse zwischen leichten Teilchen nach der Diracschen Theorie

Eine Bemerkung zur gleichnamigen Arbeit von B. Kockel

Von Peter Havas

Inhaltsübersicht

Ein kürzlich von B. Kockel aufgestellter Satz, demzufolge keine Prozesse möglich sind, in denen bei Beteiligung beliebig vieler Lichtquanten mehr als ein Elektron seinen Zustand wechselt, ist nur bei Vernachlässigung der Wechselwirkung der Elektronen richtig.

In einer kürzlich erschienenen Arbeit von B. Kockel¹⁾ wurde folgender Satz aufgestellt: „Die Übergangswahrscheinlichkeit ist gleich Null für solche Prozesse höherer Ordnung zwischen Elektronen und Lichtquanten, bei denen sich Anfangs- und Endzustand dadurch voneinander unterscheiden, daß zwei oder mehr Elektronen in einen anderen Zustand gelangt sind“²⁾.

Zum Beweis dieses Satzes wurde die Übergangswahrscheinlichkeit nach folgender Formel berechnet:

$$W = \frac{2\pi}{\hbar \Delta E_n} \left| \sum \frac{V_{12} V_{23} V_{34} \dots V_{n-1, n}}{(E_1 - E_2)(E_1 - E_3) \dots (E_1 - E_{n-1})} \right|^2. \quad (1)$$

Dabei wurde das Matricelement

$$V_{k, k+1} = e \sqrt{\frac{c \hbar}{g V}} \{ a_{\vec{p}, \sigma', \lambda'}^* (\vec{\alpha} \cdot \vec{e}) a_{\vec{p} \pm \vec{q}, \sigma'', \lambda''} \} \quad (2)$$

gesetzt, was dem Übergang eines durch eine ebene Welle beschriebenen Elektrons in einen anderen Zustand unter Aufnahme des Impulses \vec{q} eines absorbierten Lichtquants oder unter Abgabe des Impulses \vec{q} eines emittierten Lichtquants entspricht. Für den Fall, daß sich Anfangs- und Endzustand dadurch unterscheiden, daß zwei Elektronen in einen anderen Zustand gelangt sind, wurde nun die Zustandsänderung untersucht, die sich ergibt, wenn z. B. das eine Teilchen n Quanten, das andere m Quanten emittiert, bzw. absorbiert. Da die Summation in (1) sich über alle Zwischenzustände erstreckt, wurden die obigen zwei Reihen von Zustandsänderungen „ineinandergeschachtelt“, wobei sich ergab, daß der Beitrag zu (1) einer solchen Kombination von Übergangsreihen gleich Null ist. Daraus wurde auf den anfangs zitierten Satz geschlossen.

Dieses Ergebnis ist aber unvollständig. Wenn sich in der Diracschen Theorie in einer bestimmten Ordnung eine Übergangswahrscheinlichkeit Null ergibt,

¹⁾ B. Kockel, Ann. Physik (6) **4**, 279 (1949).

²⁾ B. Kockel, a. a. O., S. 286.

dann müssen eben die höheren Ordnungen untersucht werden. Dies ist im obigen Falle sehr einfach möglich. Die Matrixelemente (2) sind nämlich nicht die einzigen. Man muß auch die der Wechselwirkung zwischen den Elektronen entsprechenden berücksichtigen. Diese Wechselwirkung besteht aus zwei Teilen³⁾: einem Coulombschen, dessen Matrixelement (in der Kockelschen Bezeichnungsweise)

$$V_C = \frac{4\pi\hbar^2 e^2}{|\Delta\vec{p}|^2} \left(a_{p_1', \sigma_1', \lambda_1'}^* a_{p_1', \sigma_1', \lambda_1'} \right) \left(a_{p_2', \sigma_2', \lambda_2'}^* a_{p_2', \sigma_2', \lambda_2'} \right) \quad (3)$$

ist, und einer Emission und Reabsorption von „virtuellen“ Lichtquanten. Im Falle der einfachen Streuung wird die gesamte Wechselwirkung durch das Möllersche Matrixelement

$$V_M = \frac{4\pi\hbar^2 e^2}{|\Delta\vec{p}|^2 - \left(\frac{\Delta E}{c}\right)^2} \left[\left(a_{p_1', \sigma_1', \lambda_1'}^* a_{p_1', \sigma_1', \lambda_1'} \right) \left(a_{p_2', \sigma_2', \lambda_2'}^* a_{p_2', \sigma_2', \lambda_2'} \right) - \left(a_{p_1', \sigma_1', \lambda_1'}^* \vec{\alpha} a_{p_1', \sigma_1', \lambda_1'} \right) \left(a_{p_2', \sigma_2', \lambda_2'}^* \vec{\alpha} a_{p_2', \sigma_2', \lambda_2'} \right) \right] \quad (4)$$

beschrieben⁴⁾. Der Fall der Elektronenstreuung mit gleichzeitiger Emission eines Lichtquants wurde ebenfalls zuerst von Möller behandelt⁵⁾, der auch hier eine nichtverschwindende Übergangswahrscheinlichkeit erhielt. In zwei Arbeiten des Verfassers wurde der allgemeine Fall der Wechselwirkung beliebig vieler Lichtquanten mit zwei⁶⁾ und mit beliebig vielen⁷⁾ Elektronen behandelt. Die „Ineinandererschachtelung“ der verschiedenen Reihen von Zustandsänderungen wird durch die Beteiligung der virtuellen Quanten und der Coulombschen Wechselwirkung wesentlich kompliziert; die Summation kann aber ausgeführt werden und ergibt ein einfaches, von Null verschiedenes Resultat⁸⁾. Die virtuellen Quanten treten nicht mehr explizit auf; ihr Effekt ist einfach, daß überall das Coulombsche Matrixelement (3) durch das Möllersche (4) zu ersetzen ist, wobei nun ΔE analog zu $\Delta\vec{p}$ aus den Anfangs- und Endenergien der beteiligten Lichtquanten und Elektronen zu bilden ist.

Die Übergangswahrscheinlichkeit für Prozesse höherer Ordnung mit gleichzeitiger Zustandsänderung mehrerer Elektronen verschwindet also nicht, falls nur die Wechselwirkung der Teilchen konsequent in Rechnung gestellt wird.

³⁾ Vgl. W. Heitler, *The Quantum Theory of Radiation* (Oxford 1936), S. 98.

⁴⁾ C. Möller, *Z. Physik* **70**, 786 (1931) u. *Ann. Physik* (5) **14**, 531 (1932). Ableitung mittels der Dirac-Fermischen Methode: H. Bethe u. E. Fermi, *Z. Physik* **77**, 296 (1932).

⁵⁾ C. Möller, *Proc. Roy. Soc. London A* **152**, 481 (1935). Ableitung mittels der Dirac-Fermischen Methode: K. Umeda, *J. Fac. Sc. Hokkaido Imp. Univ. (II)* **1**, 295 (1937).

⁶⁾ P. Havas, *Physic. Rev.* **66**, 69 (1944).

⁷⁾ P. Havas, *Physic. Rev.* **68**, 214 (1945).

⁸⁾ Gl. (15) von ⁶⁾ und (15), (21) und (29) von ⁷⁾. Im Laufe der Rechnung treten den Kockelschen sehr ähnliche Reihen auf. Der wesentliche Unterschied ist, daß diese Reihen nun infolge der Wechselwirkung zwischen den Teilchen bei einem Zwischenzustand abbrechen, der nicht dieselbe Energie wie der Anfangszustand besitzt. In ¹⁾ hingegen enden die Reihen mit derselben Energie, was das Verschwinden der Summe zur Folge hat.

Bethlehem, Pennsylvania, USA., Department of Physics, Lehigh University.

(Bei der Redaktion eingegangen am 5. März 1950.)

Bemerkungen zu der vorstehenden Arbeit

Von B. Kockel

In der Arbeit mit dem obigen Titel habe ich mich von der Einleitung an auf solche Prozesse beschränkt, in denen die Elektronenübergänge durch Absorption oder Emission von Lichtquanten geschehen. Bereits die einfache Coulomb-Streuung zweier Elektronen aneinander ist natürlich ein Fall, in dem „zwei Elektronen in einen anderen Zustand“ gelangen. P. Havas teilt nun Resultate mit, zu denen man gelangt, wenn man auch noch solche Matrixelemente der Störungstheorie mit berücksichtigt, die Coulombsche Wechselwirkung und Wechselwirkung durch Emission und Reabsorption virtueller Lichtquanten (Coulombsche Wechselwirkung und Retardierung) bedeuten.

Untersuchungen über die spektrale Empfindlichkeit von Selen-Photoelementen

Von Hans-Georg Sanner

(Mit 1 Abbildung)

Inhaltsübersicht

Die nach neuartigem Verfahren mit einer Cd-Zwischenschicht versehenen Selen-Photoelemente der Firma Dr. B. Lange weisen eine erhöhte Infrarotempfindlichkeit auf. Es wird gezeigt, daß dies einen charakteristischen Unterschied der Kurven der relativen spektralen Empfindlichkeit dieser Elemente gegenüber den Kurven der gewöhnlichen Selen-Photoelemente bedingt. Weiterhin wird gezeigt, daß Unregelmäßigkeiten der Oberfläche bewirken können, daß ein Lichtbündel, welches an verschiedenen Stellen auftrifft, unterschiedliche Photoströme erzeugt. Schließlich werden Ermüdungserscheinungen untersucht, die bei Bestrahlung der Photoelemente mit charakteristischen Rotwellenlängen auftreten.

Die Kurven der relativen spektralen Empfindlichkeit

Die spektrale Empfindlichkeit ist bestimmt als die Abhängigkeit des Photostromes von der Wellenlänge eingestrahltten monochromatischen Lichtes, wenn

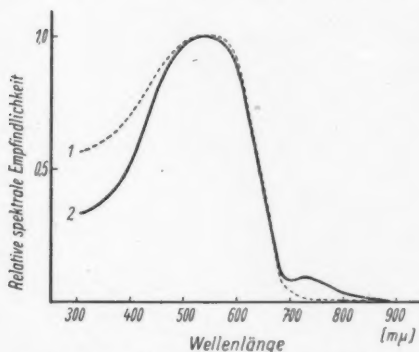


Abb. 1. Die relative spektrale Empfindlichkeit.
1 Photoelement II, 2 Photoelement 14/21 mit
Cd-Zwischenschicht

Wir erkennen, daß oberhalb einer Wellenlänge von 700 mμ die Kurve nochmals ansteigt und bei 725 mμ ein zweites kleineres Maximum aufweist, worauf die

dieses ein äquiennergetisches Spektrum besitzt. Setzt man das Maximum der so dargestellten Kurve 1 und reduziert die übrigen Ordinaten in entsprechendem Verhältnis, so erhält man die Kurve der relativen spektralen Empfindlichkeit φ_λ .

Diese wurden für ein Selen-Photoelement in der bisher üblichen Ausführung (Element II, 28 mm Durchm.) und für ein mit Cd-Zwischenschicht versehenes (Element 14/21, 28 mm Durchm.) aufgenommen. Die beiden Kurven sind in Abb. 1 dargestellt.

Kurve 1 bietet keinerlei Besonderheiten. Anders hingegen ist das Aussehen der Kurve 2.

enen
arot-
inter-
mente
erhin
B ein
hoto-
e bei
reten.

hoto-
wenn
Spek-
das
ellten
origen
n Ver-
Kurve
pfind-

Selen-
er üb-
nt II,
in mit
ehenes
rchm.)
Kurven

ei Be-
ngegen
rve 2.
chmals
uf die

geste
erfol
nach
Lich

M
der
men
von
Phot

D
empl
einer
norm
erwä
liche

D
Halb
beset
atom
therm
(Defe
(Das
leitur

E
einer
eine
platte
des
Meta
eine

N
maße
 10^{-8}
elekt
aufne
Span
rands
richtu
die
(Fluß

In
die V
strahl
Elekt

1)
2)

An

gesteigerte Empfindlichkeit für infrarotes Licht zurückzuführen ist. Dafür erfolgt aber bei dem neuen Photoelement ein stärkerer Abfall der Empfindlichkeit nach dem kurzwelligen Bereich. Eine Steigerung der Empfindlichkeit im roten Licht scheint also mit einem Verlust im violetten verbunden zu sein.

Messungen an weiteren Photoelementen mit Cd-Zwischenschicht ergaben, daß der Verlauf der spektralen Empfindlichkeit mit einem 2. Maximum für diese Elemente charakteristisch ist. Das 2. Maximum liegt bis auf geringe Verschiebungen von einigen $m\mu$ bei $725 m\mu$. Die Ordinate des Maximums variiert bei den einzelnen Photoelementen in stärkerem Maße.

Über die Ursachen der gesteigerten Infrarotempfindlichkeit

Der Verlauf der Kurve 2 legt die Vermutung nahe, daß die erhöhte Infrarotempfindlichkeit der Photoelemente mit Cd-Zwischenschicht durch Überlagerung einer glockenförmigen Kurve im Bereich von $700-1000 m\mu$ über eine annähernd normal verlaufende φ_λ -Kurve entsteht. (Gleichzeitig tritt aber noch — wie bereits erwähnt — ein stärkerer Abfall im Violetten auf.) Wie läßt sich nun die zusätzliche Empfindlichkeit im Infraroten erklären?

Das Selen ist bekanntlich ein Störstellenhalbleiter. Nach dem Wilsonschen Halbleitermodell befinden sich oberhalb des beim absoluten Nullpunkt vollbesetzten Energiebandes Störstellenniveaus (vermutlich fehlgeordnete Selenatome), die als Elektronenfänger dienen. Die Elektronen, die sich an ihnen durch thermische Einwirkung anlagern, lassen im Energieband „Elektronenlöcher“ (Defektelektronen) zurück, welche die Dunkelleitfähigkeit des Selen bewirken. (Das Vorzeichen des Halleffektes begründet die Annahme der Defektelektronenleitung.)

Ein gewöhnliches Selen-Photoelement ist folgendermaßen aufgebaut: Auf einer Metallgrundplatte ist eine Selschicht aufgebracht, auf dieser befindet sich eine dünne Edelmetallschicht. Die Stromabnahme erfolgt an der Metallgrundplatte, sowie an der lichtdurchlässigen Edelmetallschicht. Die Wirksamkeit des Photoelementes beruht auf der sog. „Sperrschicht“, die sich an der Grenze Metall-Halbleiter bildet. In unbelichtetem Zustand hat diese „Sperrschicht“ eine Gleichrichterwirkung zur Folge (Selsperrschichtgleichrichter).

Nach der Theorie von Schottky¹⁾ ist die Gleichrichterwirkung kurz folgendermaßen zu erklären: Bei Berührung von Metall und Halbleiter tritt in einer etwa 10^{-5} cm dicken Randschicht des Halbleiters eine völlige Verarmung an Defektelektronen sowie eine vollständige Ionisierung der dort befindlichen elektronenaufnehmenden Störstellen auf. („Erschöpfungsrandschicht“). Legt man eine Spannung an, so wird, wenn das Metall positiv ist, die Sperrschicht (Erschöpfungsrandschicht) verbreitert, und infolgedessen nimmt der Widerstand zu (Sperrichtung). Liegt das positive Potential am Halbleiter, so wird die Randschicht durch die Defektelektronen des Halbleiters aufgefüllt. Der Widerstand nimmt ab. (Flußrichtung.)

In Weiterführung dieser Vorstellung kann man sich im Anschluß an Mott²⁾ die Wirkung eines Selen-Photoelementes etwa so vorstellen, daß die Lichteinstrahlung ebenfalls einen Abbau der Sperrschicht bewirkt. Die Lichtquanten lösen Elektronen aus dem obersten — in der Halbleiter-Randschicht durch Defektelek-

¹⁾ Schottky, Z. Physik **118** (1941).

²⁾ Mott, Proc. Roy. Soc. London (Ser. A) **171**, 281 (1939).

tronenverarmung vollbesetzten — Energieband aus. Auf diese Weise verschwindet also die Sperrschicht, was das Auftreten des Photostromes zur Folge hat.

Als Elektronenfänger stehen das nächste, für einen Halbleiter praktisch leere Leitungsband, sowie zwischen den beiden Bändern gelegene, durch thermisch ausgelöste Elektronen noch nicht besetzte Störniveaus zur Verfügung. Der Abstand des obersten Randes des unteren Energiebandes bis zum tiefsten nicht besetzten Störniveau bedingt die langwellige Grenze dieses inneren photoelektrischen Effektes. Es ist zu vermuten, daß bei einem Photoelement, welches in der Halbleiterrandschicht in größerer Zahl solche Störniveaus enthält, eine erhöhte Empfindlichkeit für langwelliges Licht auftritt.

Die neuartigen Photoelemente der Firma Dr. B. Lange werden so hergestellt, daß auf die Selen-schicht metallisches Cadmium kathodisch aufgestäubt wird und darüber eine Edelmetallschicht als Ableitungselektrode. Es ist möglich, daß auf diese Weise Cd-Atome (bzw. Moleküle einer Cadmiumselenverbindung) sich als gitterfremde Störstellen in die Sperrschicht einlagern, deren Niveaus wie oben angegeben liegen. So ist eine Sensibilisierung des Photoelementes für rotes Licht möglich. In genauer Analogie zu der Sensibilisierung der photographischen Schichten tritt an der Cd-Zwischenschicht eine Absorption im kurzwelligen Gebiet auf, wodurch der stärkere Abfall der Kurve im Violetten bedingt ist.

Die gleiche Erscheinung hat bereits Görlich³⁾ an Selen-Photoelementen gefunden. Durch Aufbringung von Cd-Metall auf die Selenoberfläche erhielt er ein zweites Maximum der spektralen Empfindlichkeitsverteilung bei 710 m μ . Die Lage des 2. Maximums ist vermutlich durch die physikalische Natur der Störstellen, seine Höhe durch ihre Zahl bedingt. Es müßte möglich sein, durch Einbau passend gelegener Störstellen eine Sensibilisierung der Photoelemente für gewisse Spektralbereiche vorzunehmen.

Einige weitere Untersuchungen an Selen-Photoelementen

Bei der Durchführung der Messungen wurde die Frage aufgeworfen, ob der von einem Photoelement erzeugte Photostrom unabhängig davon ist, wie sich das einfallende Lichtbündel über die Oberfläche verteilt. Bei der vollständigen Ausleuchtung der Oberfläche, wie sie bei der praktischen Anwendung vielfach benutzt wird, spielt diese Frage ja keine Rolle. Will man aber die gesamte Energie eines Lichtbündels bestimmen, so kann dieses auf einen mehr oder minder großen Bereich der Oberfläche auftreten. Ergibt sich dann stets der gleiche Photostrom?

Um dies zu untersuchen, wurde eine punktförmige Lichtquelle mit einer Linse auf die Oberfläche einzelner Photoelemente abgebildet und diese mit dem Lichtpunkt abgetastet. Hierbei zeigte sich, daß bei verschiedenen Photoelementen die einzelnen Oberflächenbezirke keineswegs gleichartig reagierten. Es traten an verschiedenen Stellen gewisse „Empfindlichkeitsinseln“ auf, d. h. kleine Bereiche, zumeist nur wenige Quadratmillimeter der Oberfläche, an denen der Photostrom plötzlich um ein Vielfaches (gelegentlich bis auf das 20- oder mehrfache) anstieg. Oftmals waren diese Stellen durch etwas eigenartige Färbung von der Umgebung zu unterscheiden. Bei anderen Photoelementen zeigten sich diese Erscheinungen in gehäufter Maße am Rand, d. h. neben dem auf die Metalldeckschicht aufgesetzten Metallring, der als eine Ableitung der Photostroms diente. Sofern hier keine auffällige Veränderung der Oberfläche zu erkennen war, konnte man an-

³⁾ Görlich, Z. Physik **112** (1939).

nehmen, daß dies durch schlechten Kontakt des Metallrings bedingt war⁴). Es ist also bei der Verwendung von Photoelementen unbedingt auf guten Kontakt der Abnahmeelektroden mit der Edelmetalldeckschicht zu achten.

Ferner legen die beobachteten Oberflächenunregelmäßigkeiten es nahe, das Element stets so weit wie möglich auszuleuchten. Verschiedene Meßergebnisse mit dem gleichen Photoelement können unter Umständen darauf zurückzuführen sein, daß bei der Messung die Ausleuchtung in verschiedenem Maße erfolgte, wobei dann einmal eine vom Licht getroffene empfindlichere Randstelle eine Erhöhung des Photostroms zur Folge hatte.

Eine weitere Eigenschaft der Selen-Photoelemente — und zwar vornehmlich des neuen Typs — soll hier noch erwähnt werden. Die Elemente zeigten bei Bestrahlung mit rotem Licht der in der Umgebung von $700\text{ m}\mu$ gelegenen Wellenlängen auffallende „Ermüdungserscheinungen“. Während für die übrigen Wellenlängen der Photostrom selbst bei langer Lichteinstrahlung nahezu konstant blieb, ergab sich bei den angegebenen Wellenlängen ein zunächst rascher, dann langsamer werdender Abfall des Photostroms. Er betrug in der 1. Minute etwa 3–5% des ursprünglichen Wertes. Ferner ist der Abfall am stärksten, wenn man ein lange Zeit im Dunkeln ausgeruhtes Element bestrahlt. Verdunkelt man das Photoelement und belichtet kurz darauf erneut, so erhält man denselben Photostrom, der vor dem Abdunkeln herrschte. Verdunkelt man jedoch längere Zeit (je nach den Typen $\frac{1}{2}$ –3 Minuten), so erholt sich das Photoelement wieder, und man bekommt einen Photostrom, der höchstens um 1% niedriger liegt, als der zu Beginn der ersten Messung erzielte. Dieser ursprüngliche Photostrom tritt jedoch erst wieder auf, wenn sich das Element einige Stunden im Dunkeln erholt hat.

Auffällig ist, daß der charakteristische Wellenlängenbereich für die neuen Photoelemente mit Cd-Zwischenschicht gerade im Minimum der Kurve der spektralen Empfindlichkeit liegt (bei $700\text{ m}\mu$).

Den Mitarbeitern des Astrophysikalischen Observatoriums Potsdam und seinem Leiter, Herrn Prof. Dr. Kienle, bin ich für die Unterstützung bei der experimentellen Durchführung meiner Arbeit zu Dank verpflichtet. Besonders möchte ich an dieser Stelle Herrn Prof. Dr. Wempe für seine wertvollen Ratschläge und tatkräftige Förderung danken.

⁴) An sich wäre — von den Kontaktstörungen abgesehen — wegen des Widerstandes der Metalldeckschicht (Größenordnungsmäßig einige 100Ω) eine Zunahme des Photostromes von der Mitte der Oberfläche zum Rande hin zu erwarten. Wegen der Inhomogenität der Metalldeckschicht ist diese Erscheinung selten in reiner Form ausgeprägt.

Potsdam, Astrophysikalisches Institut.

(Bei der Redaktion eingegangen am 1. November 1949.)

Zur Frage intermediärer Statistiken

Von Henning Müller

Inhaltsübersicht

Es wird die Möglichkeit einer Anwendung intermediärer Statistiken auf Modelle diskutiert, welche aus verschiedenen unterscheidbaren Sorten Bosescher bzw. Fermischer Teilchen bestehen. Im Grenzfall, wenn alle Teilchen unterscheidbar sind, ergibt sich natürlich die Boltzmannsche Statistik.

Gentile¹⁾ hat eine allgemeine Funktion für die Verteilung der Elemente einer abgeschlossenen Gesamtheit gleichartiger, nicht unterscheidbarer Teilchen auf die Energieskala angegeben, die je nach dem Wert eines Parameters zur Fermischen, Boseschen oder auch zu einer sogenannten intermediären Verteilungsfunktion wird; dieser Parameter ist dabei gleich der maximal zulässigen Besetzungszahl einer Zelle des Phasenraumes, d. h. als der eine Grenzfall folgt die Fermi-Statistik für den Wert 1, und als der andere Grenzfall folgt die Bose-Statistik für den Wert N bzw. ∞ dieses Parameters (N ist die Gesamtteilchenzahl). Sommerfeld²⁾ hat zur Frage der intermediären Statistiken (im Sinne Gentiles) bemerkt, daß sie zur Beschreibung gleichartiger, nicht unterscheidbarer Teilchen nicht herangezogen werden können, da die Wellenfunktionen solcher Systeme stets nur symmetrisch oder antisymmetrisch in den Teilchenkoordinaten sind, entsprechend den Boseschen oder Fermischen Verteilungsfunktionen.

Im Anschluß an die Arbeiten von Gentile hat Schubert³⁾ zunächst gezeigt, daß es gleichgültig ist, ob man in der erweiterten Boseschen Verteilungsfunktion N oder ∞ als maximale Besetzungszahl einer Elementarzelle annimmt, und zwar deshalb, weil der Fehler, den man durch Anwenden der Stirlingschen Formel in der auch von Gentile benutzten Form $n! \sim n^n e^{-n}$ macht, stärker ins Gewicht fällt, als die scheinbare Verbesserung, die durch die Annahme endlicher maximaler Besetzungszahlen erreicht wird. Weiter hat Schubert mit Hilfe der Darwin-Fowlerschen Sattelpunktmethode einen asymptotischen Ausdruck für die Bosesche Verteilungsfunktion hergeleitet, der sich für große Teilchenzahlen auf ein einziges Glied reduziert, welches dem gleicht, zu dem man auch durch Anwenden der Stirlingschen Formel gelangt.

In diesem Zusammenhang soll die Frage gestellt werden, ob es überhaupt ein — denkbares — statistisches Modell gibt, das durch intermediäre Verteilungsfunktionen — im Gentileschen oder auch in anderem Sinne — beschreibbar ist.

Die Elemente einer Gesamtheit nichtunterscheidbarer Teilchen gehorchen nur entweder der Bose- oder der Fermi-Statistik; dies folgt allein aus der Forderung,

¹⁾ Gentile, *Nuovo Cimento Ann.* **17**, 10 (1940).

²⁾ Sommerfeld, *Ber. dtsch. chem. Ges.* **75**, 1988 (1942).

³⁾ Schubert, *Z. Naturforschg.* **1**, 114 (1946).

daß die Vertauschung zweier Teilchen nicht zu einem neuen Mikrozustand führen soll. Die Frage nach dem Verhalten von Systemen gleichartiger, individuierbarer Teilchen ist wiederum durch die Boltzmann-Statistik bereits gelöst. Es erscheint deshalb zweckmässig, Gesamtheiten zu betrachten, deren Elemente zum Teil unterscheidbar und zum Teil nicht unterscheidbar sind. Einerseits ist ein plausibler physikalischer Grund für die Annahme einer zwischen 1 und N liegenden maximalen Besetzungszahl — dies würde einer intermediären Verteilungsfunktion im Sinne Gentiles entsprechen — offenbar schwer zu finden; andererseits aber erscheint es unabhängig von dem Vorgehen Gentiles möglich zu sein, vernünftige intermediäre Verteilungsfunktionen auf andere Weise zu erzeugen; und zwar durch Verallgemeinerung einer Hypothese von Brillouin⁴⁾, nach welcher jedem Teilchen ein bestimmtes Eigenvolumen im Phasenraum zugeordnet ist.

In welcher Weise können sich unter der Annahme eines solchen Eigenvolumens n_i — zunächst unterscheidbare — Teilchen auf die z_i Zellen der i -ten Energiezone verteilen? Die Volumeneinheit des Phasenraumes sei durch die Größe einer Elementarzelle gegeben; das Eigenvolumen eines Teilchens sei b_i mal dieser Volumeneinheit. Dem ersten der zu verteilenden n_i Teilchen steht danach ein Volumen z_i zur Verfügung, dem zweiten ein solches $z_i - b_i$ usw. Allgemein kann sich das k -te Teilchen der i -ten Zone seinen Platz in einem Raum von der Größe

$$V_{ik} = z_i - (k-1)b_i \quad (1)$$

ausuchen. Für die — entsprechend normierte — Wahrscheinlichkeit einer beliebigen Verteilung innerhalb der i -ten Zone folgt danach

$$W_i = \prod_k V_{ik} \quad (2)$$

und die Wahrscheinlichkeit eines Makrozustandes ergibt sich zu

$$W = N! \prod_i \frac{W_i}{n_i!}, \quad (3)$$

wobei $N!$ und $n_i!$ dem Umstand Rechnung tragen, daß die Vertauschung zweier Teilchen aus verschiedenen Zonen zu einem neuen Mikrozustand führt, während dies bei der Vertauschung zweier Teilchen der gleichen Zone nicht der Fall ist; denn innerhalb einer Zone ist ja den verschiedenen Verteilungsmöglichkeiten bereits durch das Produkt (2) Rechnung getragen. Für den Gleichgewichtszustand folgt unter der Voraussetzung konstanter Gesamtteilchenzahl und konstanter Gesamtenergie unter der (wesentlichen) Annahme $z_i \gg 1$ und unter Verwendung der Stirlingschen Formel die allgemeine Verteilungsfunktion

$$n_i = \frac{z_i}{e^{\alpha + \beta E_i} + b_i} = \frac{z_i \varrho(T) \exp\left(-\frac{E_i}{kT}\right)}{1 + b_i \varrho(T) \exp\left(-\frac{E_i}{kT}\right)} \quad (4)$$

(mit $\varrho \equiv e^{-\alpha}$ und $\beta = \frac{1}{kT}$), die je nachdem, ob der Parameter b_i (das Eigenvolumen der Teilchen) die Werte 0, 1 oder -1 annimmt, dem Boltzmannschen, Fermischen oder Boseschen statistischen Modell entspricht, d. h. je nach dem Wert von b_i beschreibt die Verteilungsfunktion eine Gesamtheit Boltzmannscher bzw. Fermischer bzw. Bosescher Teilchen.

⁴⁾ Brillouin, Quantenstatistik, Berlin 1931.

Der Parameter b_i muß nun offenbar dann andere als diese 3 konstanten Werte 0, 1, -1 annehmen, wenn es sich um eine Gesamtheit (I) handelt, die aus mehreren Sorten Teilchen gebildet wird, derart, daß sich die Teilchen verschiedener Sorten voneinander unterscheiden, die Teilchen einer Sorte unter sich aber nicht unterscheidbar sind. Die Frage ist die: Kann man einer solchen Gesamtheit (I) stets eine Gesamtheit (II) einheitlicher unterscheidbarer Teilchen zuordnen, welche phänomenologisch das gleiche physikalische Verhalten zeigt, und ist insbesondere eine solche Gesamtheit einheitlicher Teilchen durch eine Verteilungsfunktion der Form (4) beschreibbar? Die reale Existenz einer Gesamtheit (II) ist ausgeschlossen, da bei Gesamtheiten einheitlicher Teilchen aus theoretischen Gründen, wie gesagt, nur die Eigenvolumina 1, -1, 0 realisiert sind. Es kann sich bei der Annahme einer solchen Gesamtheit also nur um ein methodisches Hilfsmittel handeln, durch dessen Verwendung gezeigt werden soll, welche Arten intermediärer Verteilungsfunktionen — nicht im Sinne Gentiles (das hieße in bezug auf die maximale Besetzungszahl der Elementarzelle des Phasenraumes), sondern im Sinne variablen Eigenvolumens — durch Gesamtheiten (I) realisiert werden können.

Ein besonders einfacher Fall liegt sicher dann vor, wenn die Anteile der einzelnen Sorten an der Gesamtzahl der Teilchen für jede Energiezone die gleichen sind. Im allgemeinen wird dies aber nicht der Fall sein, und zwar deshalb nicht, weil die Verteilungsfunktionen zweier abgeschlossener Bose- oder Fermi-Gesamtheiten gleichen Volumens, gleicher Temperatur aber verschiedener Gesamtteilchenzahl sich nicht nur — wie es bei Boltzmannschen Gesamtheiten der Fall wäre — um einen konstanten Faktor unterscheiden, sondern grundsätzlich verschiedenes Verhalten zeigen. Das heißt aber nichts anderes, als daß die Anteile jeder Teilchensorte für die einzelnen Zonen verschieden, mithin energieabhängig sind.

Wie sieht das Eigenvolumen b_i als Funktion der Sortenanteile in der i -ten Energiezone aus? Es handele sich um eine Gesamtheit, welche aus m Fermi-Sorten und n Bose-Sorten gebildet wird, und es seien:

x_{ki} das Verhältnis der in der i -ten Energiezone befindlichen Fermiteilchen der Sorte k zur Gesamtzahl n_i der Teilchen, welche sich in dieser Zone befinden, und

y_{li} das Verhältnis der in der i -ten Energiezone befindlichen Boseteilchen der Sorte l zur Gesamtteilchenzahl dieser Zone.

$1 - b_i$ bezeichnet das Volumen, welches in einer Zelle, die schon ein Teilchen beherbergt, noch frei ist, anders ausgedrückt: $1 - b_i$ ist die Wahrscheinlichkeit dafür, daß eine Zelle der i -ten Zone, die bereits mit einem Teilchen besetzt ist, ein weiteres Teilchen aufnehmen kann. Sie ergibt sich im angenommenen Fall zu:

$$1 - b_i = 2 \sum_l y_{li}^2 + 2 \sum_{k,l} x_{ki} y_{li} + \sum_{k,j} \delta_{kj} x_{ki} x_{ji} + \sum_{l,r} \delta_{lr} y_{li} y_{ri}. \quad (5)$$

Denn: ist in der Zelle ursprünglich ein Fermiteilchen, so hindert dieses die Aufnahme irgendeines anderen Teilchens (mit Ausnahme eines solchen der gleichen Sorte) nicht, und gleiches gilt für den Fall, daß die Teilchen miteinander vertauscht sind; das entspricht dem Ausdruck:

$$2 \sum_{k,l} x_{ki} y_{li} + \sum_{k,j} \delta_{kj} x_{ki} x_{ji}.$$

Ein Boseteilchen hindert nicht die Aufnahme irgendeines anderen Boseteilchens; gehört das andere der gleichen Sorte an, so ist dieser Fall doppelt zu

zählen, da die beiden Teilchen — das, welches zuerst in der Zelle war, und das, welches hinzukommt — ihre Rollen ja auch wiederum vertauschen können; diese Möglichkeiten entsprechen den Gliedern

$$2 \sum_i y_{ii}^2 + \sum_{r,l} \delta_{rl} y_{li} y_{ri}.$$

Fassen wir die Glieder auf der rechten Seite von (5) in geeigneter Weise unter Berücksichtigung von

$$1 = \sum_k x_{ki}^2 + \sum_l y_{li}^2 + 2 \sum_{k,l} x_{ki} y_{li} + \sum_{k,j} \delta_{kj} x_{ki} x_{ji} + \sum_{l,r} \delta_{lr} y_{li} y_{ri} \quad (6)$$

zusammen, so haben wir als Ergebnis

$$b_i = \sum_k x_{ki}^2 - \sum_l y_{li}^2. \quad (7)$$

Es ist also möglich, eine Gesamtheit (I), welche aus mehreren Sorten nicht unterscheidbarer Teilchen besteht, so zu beschreiben, als ob sie eine einheitliche Gesamtheit (II) gleichartiger Teilchen wäre. Das Eigenvolumen b_i eines Teilchens von (II) ergibt sich als Funktion der jeweiligen Sortenanteile.

Als einfachstes Beispiel der Fall, daß die Gesamtheit (I) aus p Sorten Fermi- bzw. p Sorten Bose-Teilchen bestehen soll, wobei jede Sorte durch die gleiche Teilchenzahl vertreten sei. Das Eigenvolumen b_i hängt nach (7) von p gemäß

$$b_i = \frac{1}{p} \quad \text{bzw.} \quad b_i = -\frac{1}{p} \quad (8)$$

ab. Der Entartungsfaktor $\varrho(T)$ wird — bei einem einheitlichen Gas — durch Summation bzw. Integration von (4) über alle Zonen unter Berücksichtigung von

$$\sum_i n_i = N$$

bestimmt. Es entsteht folgender Zusammenhang⁵⁾:

$$\left(\frac{T_0}{T}\right)^{3/2} = \sum_r \varrho^r(T) (\pm 1)^{r-1} r^{-3/2}, \quad (9)$$

wobei rechts das (+)-Zeichen dem Bose-Fall und das (—)-Zeichen dem Fermi-Fall entspricht; T_0 ist die sog. Entartungstemperatur. In (9) muß offensichtlich $\varrho \leq 1$ sein; für $\varrho \gg 1$ gilt (nur für das Fermische Modell; im Boseschen Fall ist stets $\varrho \leq 1$) die asymptotische Formel

$$\left(\frac{T_0}{T}\right)^{3/2} \sim \frac{4}{3\sqrt{\pi}} \ln^{3/2} \varrho(T). \quad (10)$$

Die Entartungstemperatur T_0 , die bei einheitlichen Gesamtheiten gemäß

$$T_0 = \frac{(Nv)^{2/3}}{V^{1/3} 2\pi m k}. \quad (11)$$

v ist das Volumen einer Elementarzelle, V das Systemvolumen, m die Teilchenmasse und k die Boltzmannsche Konstante) proportional $N^{2/3}$ ist, ist bei dem jetzt ins Auge gefaßten Fall (gekennzeichnet durch ϱ^* und T_0^*) proportional zu

⁵⁾ Vgl. etwa Tolman, The principles of statistical mechanics, Oxford 1948.

$(N/p)^{1/2}$, da die Zahl der Teilchen einer jeden Sorte gleich N/p ist. Mithin tritt an die Stelle von (9)

$$\sum_{\tau} e^{*\tau} (T) (\pm 1)^{\tau-1} \tau^{-1/2} = \left(\frac{T_0}{T}\right)^{1/2} \frac{1}{p} = \left(\frac{T_0}{T}\right)^{1/2}, \quad (12)$$

d. h. es geht insbesondere $q^* \rightarrow 0$ für $p \rightarrow \infty$ bei festem T und für $T \rightarrow \infty$ bei festem p . Dieses Ergebnis entspricht der Erwartung, daß sich im Grenzfalle $p \rightarrow \infty$ (eigentlich $p \rightarrow N$) die Boltzmannsche Statistik ergeben muß.

Sind Gesamtheiten (I) realisierbar?

Sie müssen aus voneinander unabhängigen Teilchen gleicher Masse bestehen, die z. T. unterscheidbar und z. T. nicht unterscheidbar sind. Da zusammengesetzte Teilchen sich unabhängig von der Anzahl der Boseteilchen nur gemäß der Anzahl der Fermiteilchen, aus denen sie aufgebaut sind, in bezug auf ihr statistisches Verhalten unterscheiden, und zwar in dem Sinne, daß sie sich dann, wenn sie aus einer ungeraden Anzahl von Fermiteilchen bestehen, entsprechend der Fermi-Statistik, wenn sie aus einer geraden Anzahl von Fermi-Teilchen bestehen, entsprechend der Bose-Statistik verhalten, bilden mehrere Sorten isobarer Atome oder Moleküle bei hinreichend tiefer Temperatur eine Gesamtheit (I); denn die Elementarpartikel: Proton, Neutron und Elektron sind ja Fermi-Teilchen.

Zur Frage der gedanklichen Äquivalenz von Gesamtheiten, die (a) aus einheitlichen Teilchen und (b) aus verschiedenen Teilchensorten bestehen, ist demnach zu sagen: Wenn man — wie es hier geschah — von unterscheidbaren Teilchen ausgeht, so führt ein Eigenvolumen -1 bzw. 1 auf die gleichen Verteilungsfunktionen, wie sie Gesamtheiten nichtunterscheidbarer Teilchen entsprechen würden. Eine Gesamtheit unterscheidbarer Teilchen mit einem von -1 bzw. 1 verschiedenen Eigenvolumen kann gedeutet werden:

1. als Gesamtheit unterscheidbarer Teilchen vom Eigenvolumen b_i oder

2. als Gesamtheit mehrerer Sorten jeweils nichtunterscheidbarer Teilchen mit den Eigenvolumina -1 bzw. 1 .

Es sei noch darauf hingewiesen, daß es nicht unvernünftig erscheint, die Fragestellung, die zur Verwendung intermediärer Statistiken führte, umzukehren, also nicht vorgegebene Gesamtheiten zu beschreiben, sondern vorgegebene Verteilungsfunktionen zu realisieren. Das würde heißen, diejenige Zusammensetzung einer Gesamtheit zu finden, die einem vorgegebenen, energieabhängigen Eigenvolumen $b_i(E_i)$ entspricht.

Mainz, Institut für theoretische Physik der Universität.

(Bei der Redaktion eingegangen am 27. April 1950.)

Verantwortlich

für die Schriftleitung: Prof. Dr. Friedrich Möglich, Berlin-Buch, Lindenberger Weg 74; für den Anzeigenteil: Ernst Wöllnitz (Arbeitsgemeinschaft medizinischer Verlage G.m.b.H.), Berlin C 2, Neue Grünstraße 18, Fernruf: 52 20 97. Z. Z. gilt Anzeigenpreisliste Nr. 1 Verlag: Johann Ambrosius Barth, Leipzig C 1, Salomonstr. 18 B, Fernruf: 63 106, 63 781 (1) Paul Dünhaupt, Köthen L 256/50 Lizenz-Nr. 410 / G.N. 446

6. FOLGE, 1950, BAND 7, HEFT 7-8

UNIVERSITY
OF MICHIGAN

OCT 12 1950

PHYSICS
LIBRARY

ANNALEN DER PHYSIK

Gegründet 1799 durch F. A. C. Gren
Fortgeführt durch L. W. Gilbert, J. C. Poggendorff,
G. und E. Wiedemann, P. Drude,
W. Wien, M. Planck, E. Grüneisen

Der ganzen Reihe 442. Band

Kuratorium:

W. GERLACH, CHR. GERTHSEN, F. HUND, W. KOSSEL,
M. v. LAUE, W. MEISSNER, R. W. POHL, R. ROMPE,
R. SEELIGER, A. SOMMERFELD, W. WEIZEL

Herausgegeben von

F. MÖGLICH
Berlin



JOHANN AMBROSIOUS BARTH / VERLAG / LEIPZIG

Redaktionschluß am 20. Juli 1950

Inhalt

	Seite
R. Borchert, Berlin-Buch: Die Größe der effektiven Ionisierungsspannung in der Quecksilberhochdruckentladung. (Mit 5 Abbildungen).....	321
R. Borchert, Berlin-Buch: Die Entstehung des Kontinuums in der Quecksilberhochdruckentladung und die Größe ihrer mittleren Anregungsspannung. (Mit 1 Abbildung).....	332
Hans Wolter, Kiel: Die Minimumstrahlkennzeichnung als Mittel zur Genauigkeitssteigerung optischer Messungen und als methodisches Hilfsmittel zum Ersatz des Strahlbegriffes. (Mit 7 Abbildungen).....	341
Karl Bechert, Mainz: Ansätze zu einer nichtlinearen Elektrodynamik	369
A. Eckardt und O. Lindig, Jena: Ausleuchtung von Phosphoren durch Ultra-Schall. (Mit 4 Abbildungen).....	410
Peter Havas, Bethlehem (Pennsylvania, USA.): Prozesse zwischen leichten Teilchen nach der Diracschen Theorie. Eine Bemerkung zur gleichnamigen Arbeit von B. Kockel	419
Hans-Georg Sanner, Potsdam: Untersuchungen über die spektrale Empfindlichkeit von Selen-Photoelementen. (Mit 1 Abbildung).....	416
Henning Müller, Mainz: Zur Frage intermediärer Statistiken	420

Für diese Zeitschrift werden grundsätzlich nur Arbeiten angenommen, die vorher weder im Inland noch im Ausland veröffentlicht worden sind. Mit der Annahme des Manuskriptes und seiner Veröffentlichung geht das alleinige Recht der Vervielfältigung, Verbreitung und Übersetzung einschließlich des Rechts der Veranstaltung von Sonderdrucken auf den Verlag über. Von jeder Arbeit liefert der Verlag 75 Sonderabzüge kostenlos.

Es wird gebeten, Manuskripte an Herrn Prof. Dr. F. Möglich, (1) Berlin-Buch, Lindenberger Weg 74, einzusenden.

Die Redaktion bittet, die Manuskripte *druckfertig* einzuliefern und in den Korrekturen den bisherigen Raum des Textes nicht zu überschreiten. Die *Zeichnungen* sind in sorgfältiger Ausführung auf besonderen Blättern beizulegen.

Bei den *Zitaten* wolle man die Abkürzungen der Zeitschriftentitel nach dem Verzeichnis wählen, das jedem Jahrgang der „Physikalischen Berichte“ vorangestellt ist.

Die Zeitschrift erscheint in Bänden mit 8 Heften, die zwanglos ausgegeben werden. Die Lieferung erfolgt durch den Sortimentsbuchhandel, in Ausnahmefällen unmittelbar vom Verlag. Der Bezugspreis beträgt je Band DM. 34.— zuzügl. Postgebühren. Einzelpreis dieses Heftes DM 10.40.

Die Lieferung erfolgt bis zur Abbestellung, die nur für das Ende eines Bandes ausgesprochen werden kann.

Seite

321

332

341

369

410

413

416

420

ber
des
ng,
er-
ge

n-

k-
nd

r-
t.

n.
l.
n.

s

X